

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:39:45 ; Search time 2719.24 Seconds

(without alignments)
16699.717 Million cell updates/sec

Title: US-09-744-748-3

Perfect score: 2170 1 cttccctcgcctggatcgc.....aaaaaaaaaaaaaaaaaaaaa 2170

Sequence: 1 cttccctcgcctggatcgc.....aaaaaaaaaaaaaaaaaaaaa 2170

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_da.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_hcg_hum.*

31: em_hcg_inv.*

32: em_hcg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description

1	2139	98.6	2139	10	AB015426	AB015426 Mus muscu
2	1588.2	73.2	2156	10	AB049819	AB049819 Rattus no
3	1309.2	60.3	1705	10	AF230460	AF230460 Cricetulu
4	1206.8	55.6	3019	9	AB023021	AB023021 Homo sapi
5	1205	55.5	2501	9	HS428701	HS428701 Homo sapi
6	1162.4	53.6	80247	9	AL512406	AL512406 Human DNA
7	1049.6	48.4	71396	10	AF345993	AF345993 Rattus no
8	883	40.7	71396	5	AC100674	AC100674 Mus muscu
9	740	34.1	71396	2	AC100674	AC100674 Mus muscu
10	740	34.1	71396	2	AC100674	AC100674 Mus muscu
11	675.2	31.1	1080	5	AB035905	AB035905 xenopus 1
12	269	12.4	1567	5	AB023627	AB023627 Dario rer
13	230.4	10.6	1832	5	AB023628	AB023628 Dario rer
14	220.6	10.2	1264	5	AF288369	AF288369 Gallus ga
15	206.6	9.5	1092	5	AB035907	AB035907 xenopus 1
16	175.2	8.1	2657	5	GGU73678	GGU73678 Gallus gall
17	157	7.2	1429	10	CGU78737	CGU78737 Cricetulu
18	157	7.2	1462	10	AF090450	AF090450 Cricetulu
19	155.4	7.2	1718	10	AF090449	AF090449 Cricetulu
20	135.6	6.2	1044	5	AB035908	AB035908 Gallus ga
21	134.8	6.2	2293	4	BTAL134FTG	BTAL134FTG Bos tauru
22	134.8	6.2	12607	4	BTAL132772	BTAL132772 Bos tauru
23	134.2	6.2	1158	10	AB039126	AB039126 Mus muscu
24	133.4	6.1	1113	9	AF345884	AF345884 Salimlrl s
25	133	6.1	1119	9	AF345881	AF345881 Macaca mu
26	132.6	6.1	1170	10	AB039127	AB039127 Mus muscu
27	132.6	6.1	1170	10	AB039127	AB039127 Mus muscu
28	132.6	6.1	1615	10	AB039131	AB039131 Mus muscu
29	131.4	6.1	1170	10	AB039133	AB039133 Mus muscu
30	131	6.0	1170	10	AB039124	AB039124 Mus muscu
31	131	6.0	1170	10	AB039125	AB039125 Mus muscu
32	131	6.0	1170	10	AB039128	AB039128 Mus muscu
33	131	6.0	1170	10	AB039129	AB039129 Mus muscu
34	131	6.0	1170	10	AB039130	AB039130 Mus muscu
35	131	6.0	1170	10	AB039132	AB039132 Mus muscu
36	130.8	6.0	1302	10	AB049938	AB049938 Rattus no
37	130.8	6.0	169446	2	AC097778	AC097778 Rattus no
38	130	6.0	1613	4	BTAL132776	BTAL132776 Bos tauru
39	130	6.0	1619	4	BTAL132775	BTAL132775 Bos tauru
40	130	6.0	1644	4	BTAL132774	BTAL132774 Bos tauru
41	130	6.0	1873	4	BTAL132773	BTAL132773 Bos tauru
42	129.8	6.0	1162	9	PT14033	PT14033 Pan troglod
43	129.8	6.0	80893	2	AC024592	AC024592 Homo sapi
44	129.8	6.0	187488	2	AC010454	AC010454 Homo sapi
45	128.8	5.9	169893	2	AC021165	AC021165 Homo sapi

ALIGNMENTS

RESULT 1

AB015426

LOCUS

DEFINITION

AB015426 Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.

ACCESSION

AB015426

VERSION

AB015426.1 GI:3702718

KEYWORDS

Fut9; alpha1,3-fucosyltransferase IX.

SOURCE

Mus musculus cDNA to mRNA.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (sites)

Kudo,T., Ikehara,Y., Togayachi,A., Kaneko,M., Hirtaga,T., Sasaki,K. and Narimatsu,H.

TITLE

Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mfuc-TIX, that synthesizes the Lewis x (CD15) epitope in brain and kidney

JOURNAL

J. Biol. Chem. 273 (41), 26729-26738 (1998)

REFERENCE

98434588

AUTHORS

Kudo,T. and Narimatsu,H.

2 (bases 1 to 2139)

Direct Submission

JOURNAL Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science,
Soka University, Division of Cell Biology, 1-236 Tangi-cho,
Hachioji, Tokyo 192-8577, Japan (E-mail:tkudo@et.soka.ac.jp,
Tel:+81-426-91-2495, Fax:+81-426-91-9315)

FEATURES

source

1. 2139

/organism="Mus musculus"

/db_xref="taxon:10090"

115..1194

/gene="Fut9"

115..1194

/gene="Fut9"

/codon_start=1

/product="alpha1,3-fucosyltransferase IX"

/protein_id="BAA33522.1"

/db_xref="GI:3702719"

/translation="MTSTSGILRPFLIVCIILGCFMACLIYIKPTNSWVSPMESA
SSVLMKMFSTKIDYENETTLVWVPEFOTDLSCQAMENIOGHLPTDRLYLNK
SHAVLIHROISMDLNLPOARPPROKIMWNLSPHTPOKSGIEHLNLTLYRR
DSDIOVPYGLTVSTNPVEFVPSKEKLUWNLSPHTPOKSGIEHLNLTLYRR
GOARPEVYNDKNIIPRTSCREYLSPEBSIKHXYITEKLYNAPLAGSVPLPSREN
YENITPADSRFHYEDPNSPELAKTLKVDNKNLILSTFNRKDFVNLPRWESHA
CLADHVARHOFKSYGNLEKMFNN"

polyA site

/note="31 a nucleotides"

BASE COUNT 684 a 395 c 368 g 692 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTCCCTCGCTGGATGCGCTTGTAGTAAGTGAATACCAACATCATATGACAGCTGGA 60
DB 1 CTTTCCCTCGCTGGATGCGCTTGTAGTAAGTGAATACCAACATCATATGACAGCTGGA 60
QY 61 GTAGTTGACGACATCTCTGCTTCCATGCTATGTTCTTACACTGAAAAATATATGACA 120
DB 61 GTAGTTGACGACATCTCTGCTTCCATGCTATGTTCTTACACTGAAAAATATATGACA 120
QY 121 TCACATCCAAAGGACATCTCTGCTTCCATGCTATGTTCTTACACTGAAAAATATATGACA 180
DB 121 TCACATCCAAAGGACATCTCTGCTTCCATGCTATGTTCTTACACTGAAAAATATATGACA 180
QY 181 ATGGCATGTCTGCTCAATTAATATCAAGCCACCAAGAGCGGCTTCAGTCCAAATGAGG 240
DB 181 ATGGCATGTCTGCTCAATTAATATCAAGCCACCAAGAGCGGCTTCAGTCCAAATGAGG 240
QY 241 TCTGCAAGTTCTGCTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 300
DB 241 TCTGCAAGTTCTGCTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 300
QY 301 GAACTACCATTCGTTGGTATGGCATTTGGGACAGACCTTTGACCTTACATCTGCG 360
DB 301 GAACTACCATTCGTTGGTATGGCATTTGGGACAGACCTTTGACCTTACATCTGCG 360
QY 361 CAAGCAATGTTCAATATCCAAAGGAGGACATCTCAACAGAGCGGCTTCAGTCCAAATG 420
DB 361 CAAGCAATGTTCAATATCCAAAGGAGGACATCTCAACAGAGCGGCTTCAGTCCAAATG 420
QY 421 TCCCATGCGGTCCTGATTCACATAGACATCAGTGGAGCTGACCTAACTTACCTGAG 480
DB 421 TCCCATGCGGTCCTGATTCACATAGACATCAGTGGAGCTGACCTAACTTACCTGAG 480
QY 481 CAGGCGAGGACCCCTTTCAGAAATGATTTGGATGAATTTAGATCACCACCTGACAC 540
DB 481 CAGGCGAGGACCCCTTTCAGAAATGATTTGGATGAATTTAGATCACCACCTGACAC 540
QY 541 CCGCAAAAGAGTGACATGAACATCTGTTCAACCTGACCTTAATCTATCCCGTGAATTC 600
DB 541 CCGCAAAAGAGTGACATGAACATCTGTTCAACCTGACCTTAATCTATCCCGTGAATTC 600
QY 601 GATATCAAGTGCCTTATGCTTTGACGAGTGAGCACAATCCCTTTGTTGTAAGTG 660

DB 601 GATATCAAGTGCCTTATGCTTTGACGAGTGAGCACAATCCCTTTGTTGTAAGTG 660
QY 661 CCAAGCAAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 CCAAGCAAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 GTCAGATTAATCAAGAGCTGACAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB 721 GTCAGATTAATCAAGAGCTGACAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 781 GGAGATTAATCAAGAGCTGACAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 840
DB 781 GGAGATTAATCAAGAGCTGACAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 840
QY 841 TCATTTGAAAACGATTCACAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
DB 841 TCATTTGAAAACGATTCACAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 901 GCTGCTGACGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 GCTGCTGACGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 GCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1020
DB 961 GCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1020
QY 1021 AAGCAAGTGAACAAAACATTAAGTGTGATCTTAACTGATGAGAAAGATTTT 1080
DB 1021 AAGCAAGTGAACAAAACATTAAGTGTGATCTTAACTGATGAGAAAGATTTT 1080
QY 1081 ACTGTAACCTGACCAAGCTTTGGGATCAATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 ACTGTAACCTGACCAAGCTTTGGGATCAATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 AGGCATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
DB 1141 AGGCATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
QY 1201 CATCATTTGACAGGAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
DB 1201 CATCATTTGACAGGAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
QY 1261 AACACGCTGATTTGGGACCGCTTAAATTTCCGCGCCGCTGAGGACCAATGATAT 1320
DB 1261 AACACGCTGATTTGGGACCGCTTAAATTTCCGCGCCGCTGAGGACCAATGATAT 1320
QY 1321 TTGGTGAATTTTAAAGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
DB 1321 TTGGTGAATTTTAAAGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
QY 1381 TGATATATGATTAAGACGACATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
DB 1381 TGATATATGATTAAGACGACATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
QY 1441 TGATATATGATTAAGACGACATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
DB 1441 TGATATATGATTAAGACGACATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
QY 1501 ATCAGATCTTAACTATTTGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
DB 1501 ATCAGATCTTAACTATTTGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
QY 1561 TGCCAAAGTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
DB 1561 TGCCAAAGTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
QY 1621 AATCACTTAATTTATATCTTTGACTTAAAGTTGAACCTATGACACTGATGATAG 1680
DB 1621 AATCACTTAATTTATATCTTTGACTTAAAGTTGAACCTATGACACTGATGATAG 1680
QY 1681 TTTTACTGTTACACATTTGTAAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
DB 1681 TTTTACTGTTACACATTTGTAAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 1740

us-09-744-748-3.rge

Page 4

Page

```

OY 352 ACATCTGCCAGCAATGTTCAATATCCAAGGGTGGCATCTCAACAAGACCGCTCATTTG 411
    |||||
DB 299 ACATCTGCCAGCAAGGATGTTCAACATCAAGGGTGGCATCTCAACAAGACCGCTCAGTG 358
OY 412 TACAAATAATCCCATCGGTCCTGATTCACATAGACATCAGCTGGGATCTGACTAAC 471
    |||||
DB 359 TACAAATAATCCCATCGGTCCTGATTCACACAGACATCAGCTGGGATCTGACTAAC 418
OY 472 TTAACCTCAGACAGGCGGACCGCTTTCAGAAATGATTTGGATTTAGAGTCAACC 531
    |||||
DB 419 TTAACCTCAGACAGGCTGAGCGGCTTTCAGAAATGATTTGGATTTAGAGTCAACC 478
OY 532 ACTCACACCCCAAAAAGAGTGGCATTTGAACCTTTGCACTGACTTAATTCGC 591
    |||||
DB 479 ACTCACACCCCAAAAAGAGTGGCATTTGAGCATTTGTCACTGACTTAATTCGC 538
OY 592 CGTGAATTCAGATTCACATGCTTATGCTCTTACAGCTGAGCAAAATCCCTTTG 651
    |||||
DB 539 CGTGAATTCAGATTCACATGCTTATGCTCTTACAGCTGAGCAAAATCCCTTTG 598
OY 652 TTTGAAGTCCAGCAAGAGAGAGTGTGCTGGTGTGAGTAACTGGAACCTGAG 711
    |||||
DB 599 TTTGAAGTCCAGCAAGAGAGAGTGTGCTGGTGTGAGTAACTGGAACCTGAG 658
OY 712 CATGCCAGGGTCAAGTATTTCACAGAGCTCAGCAAGAGTATTGAATCCACACTATGCG 771
    |||||
DB 659 CATGCCAGGGTCAAGTATTTCATATATATGAGCTCAGCAAGAGTATTGAATCCACACTATGCG 718
OY 772 CAGCATTTGGAGATACGTGACGATTAATAATCTGATTTCCACCATATCTACTGTAA 831
    |||||
DB 719 CAGCATTTGGAGATACGTGATTAATAATCTGATTTCCACCATATCTACTGTAA 778
OY 832 TTTTATCTTCAATTTGAAAACATCAATTCACAAAGATTACATCAGAAAGCTCTACAT 891
    |||||
DB 779 TTTTATCTTCAATTTGAAAACATCAATTCACAAAGATTACATCAGAAAGCTCTACAT 838
OY 892 GCATTTTGGCTGGTCACTACCTGTTGCTGGTGCATCTAGGAAAACATATGAGAT 951
    |||||
DB 839 GCATTTTGGCTGGTCACTACCTGTTGCTGGTGCATCTAGGAAAACATATGAGAT 898
OY 952 TATATTTCCAGCTGATTCATCTATCATGAGAAATTTTAACTCTCCAGTATGAGTA 1011
    |||||
DB 899 TATATTTCCAGCTGATTCATCTATCATGAGAAATTTTAACTCTCCAGTATGAGTA 958
OY 1012 AATATCTCAAGGAAGTTCACAAAACAATAAGTTGATCTAGTTACTTTAACTGAGTA 1071
    |||||
DB 959 AATATCTCAAGGAAGTTCACAAAACAATAAGTTGATCTAGTTACTTTAACTGAGTA 1018
OY 1072 AAGGATTTTACTGTAAACCTTACACGGTTTGGAAATCATGCTGCTGCAATGCGAC 1131
    |||||
DB 1019 AAGGATTTTACTGTAAACCTTACACGGTTTGGAAATCATGCTGCTGCAATGCGAC 1078
OY 1132 CATGTAAAAAGCATCAGATATATAAGTCTGTTGTAATTTAGAGAAATGTTTGAAT 1191
    |||||
DB 1079 CATGTAAAAAGCATCAGATATATAAGTCTGTTGTAATTTAGAGAAATGTTTGAAT 1138
OY 1192 TAAAGTGCATCTGTCACAGT---GAAGAAGTTTAAAGACATTAATCAAGTTT 1247
    |||||
DB 1139 TAAAGTGCATCTGTCACAGT---GAAGAAGTTTAAAGACATTAATCAAGTTT 1198
OY 1248 CAGATTAAGA--AGAAACACGCTGATTTGGGACCGCTTAAATTTCTGCTCCCT 1305
    |||||
DB 1199 AAGGATTAAGAAGAAAGAACACCTGCTTCAATCAACAGTTTAAATTTGCTCCCT 1258
OY 1306 GAGGACATGATATTTGGTG--AAATTTTAAAGATCAGATTAATGACATCACTCAAT 1363
    |||||
DB 1259 GAGTAAACATGATATTTTGAAGAAATTTTAAAGCTCAGAAATTAAGAACTCACTCA 1318
OY 1364 TGGGTTTAAATTAATCCGTATATATGATTAAGACAGCAAGAAATTAATTAATTTGCA 1423
    |||||
DB 1319 TTGG--ATTAAATTAATCCGTATATATGATTAAGACAGCAAGAAATTAATTAATTTCTCA 1377

```

```

OY 1424 CTCCTATTATAAATGTTT---ACATTTTATAGTACTGTAAAGTAATTTATGA 1481
    |||||
DB 1378 CTCCTATTATAAATGTTT---ACATTTTATAGTACTGTAAAGTAATTTATGA 1437
OY 1482 TTTACTGTTTCTACATCATGATCTTTAATCTATTGCGAGATTAATAATCATATCC 1541
    |||||
DB 1438 TTTCAATGTTTCTACATCATGATCTTTAATCTATTGCGAGATTAATAATCATATCC 1497
OY 1542 TAAATATGAGGACATTTTGGCAAGTATTATAAGTATAGATTTAGCTGTAAATCC 1601
    |||||
DB 1498 TAAATATGAGGACATTTTGGCAAGTATTATAAGTATAGATTTAGCTGTAAATCC 1556
OY 1602 TAAACAAGAAGATTAATGATCACTTAATCTTAATCTTAACTTGAATCAAGTTGAACA 1661
    |||||
DB 1557 TAAACAAGAAGATTAATGATCACTTAATCTTAACTTGAATCAAGTTGAGTTG--ACA 1615
OY 1662 CTACCTGACAACGATGATGATTTTACTTG--TACCAATTTTGAAGGAGGACCTAATGCA 1719
    |||||
DB 1616 CTGTGTGATTAATGATGATCTTTTAACTGTTTACCAATTTGTAAGACAGACTGATCA 1675
OY 1720 TGTAGCAATTAAGAAGATATGACAGAA 1749
    |||||
DB 1676 TGTGTAAATTAAGAAGATATGACAGAA 1705

```

```

RESULT 4
AB023021 3019 bp mRNA linear PRI 23-JUN-1999
LOCUS Homo sapiens FUT9 mRNA for alpha-1,3-fucosyltransferase IX,
DEFINITION complete cds.
ACCESSION AB023021
VERSION AB023021.1 GI:5139692
KEYWORDS alpha-1,3-fucosyltransferase IX; FUT9.
SOURCE Homo sapiens CDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE Kaneko,M., Kudo,T., Iwasaki,H., Ikehara,Y., Nishihara,S.,
AUTHORS Nakagawa,S., Sasaki,K., Shihina,T., Inoko,H., Saitou,N. and
Narimatsu,H.
TITLE Alpha 1,3-fucosyltransferase IX (Fuc-TIX) is very highly conserved
JOURNAL between human and mouse: molecular cloning, characterization and
REFERENCE tissue distribution of human Fuc-TIX
AUTHORS FEBS Lett. 453, 237-242 (1999)
JOURNAL 2 (bases 1 to 3019)
AUTHORS kaneko,M., kudo,T. and Narimatsu,H.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-1999) Mika Kaneko, Institute of Life Science,
Soka University, Division of Cell Biology, Soka University, 1-236
Tangi-cho, Hachioji, Tokyo 192-8577, Japan
(E-mail:mika@soka.ac.jp, Tel:81-426-91-2495(ex.5132),
Fax:81-426-91-9315)

```

```

FEATURES
source
1..3019
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
295..1374
/gene="FUT9"
295..1374
/gene="FUT9"
/codon_start=1
/product="alpha-1,3-fucosyltransferase IX"
/protein_id="P0A01685.1"
/db_xref="GI:5139693"
/translation="MTSTSKGILRPFLIVCIILICFNAACLIYIKPNSWTFSPMESA
SSYLKMKNFSTKEDFNETITLWVPMFGQTEFLSCQAMFNQGLHFTDRSLYRK
SHAVLIHRDISMDLTMLPOOARPPQKIMMNLSEPTHPQKSGIEHLNLTLYRK
DSDIQVYGLIVSTNPFVEVPSKELQVYVSNMDEHARVYVYELSKSEIHTY
GOARGEVNDKNIPTISACKFVYSEPSNKHDKYITTKLYNAFLAGSVPVVLGSRN
YENTIPDSFTHEDVNSPSLALEYLVKNNKLYLYSTFNMRKDFVNLPRWESHA
CLACDHRKROEKRSVGNLEKFWN"
BASE COUNT 937 a 594 c 541 g 947 t

```

us-09-744-748-3.rge

QY	102
Db	130

Page 6

2259

QY 2061 CATTGAAATATTTTCCAAACATCATATTTAA 2096
DB 2260 CAATTAAGATTTTCTGAGCCCTAATATTTAA 2295

RESULT 5
LOCUS HSA238701 2501 bp mRNA linear PRI 21-OCT-2000
DEFINITION Homo sapiens mRNA for alpha-3-fucosyltransferase.
ACCESSION AJ238701
VERSION AJ238701.1 GI:4741566
KEYWORDS alpha-3-fucosyltransferase; FUT9 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2501)
AUTHORS Callieu-Thomas,A.; Coullin,P.; Candelier,J.J.; Balanzino,L.;
Mennesson,B.; Oriol,R. and Mollicone,R.
JOURNAL Fut9 and Fut9 genes are expressed early in human embryogenesis
MEDLINE Glycobiology 10 (8), 789-802 (2000)
REFERENCE 2 (bases 1 to 2501)
AUTHORS Mollicone,R.
TITLE Direct Submission
JOURNAL Submitted (29-APR-1999) Mollicone R., glycobiology, INSERM U504, 16
Av. Paul Vallant-Couturier, Villejuif, 94807 Cedex, FRANCE
COMMENT related sequence: AB015426.
FEATURES
source location/Qualifiers
1..2501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/tissue.type="whole embryo"
/dev_stage="40 to 60 days"
105..2466
/gene="FUT9"
105..1184
/gene="FUT9"
/codon.start=1
/product="alpha-3-fucosyltransferase"
/protein_id="CAB41890.1"
/db_xref="GI:4741567"
/db_xref="SPTREMBL:O9Y231"
translation="MTSTSKGLRPFLVCIILGCFMACLLIYIKPTNSMIFSPWESA
SSVLMKMFSTKRDYENETITLWVWPEGFDTSCQAMNIOGCHLTDRLSYLNK
SHAVLIHHRDISMDITNLPQARPREQWIMMNLSPHTTOKSIEHLPNITLYRK
DSDIOVPYGLFVSTNPFVFEVPSKEKLVWVSWNPEHAKYINLSISIEIHTY
GQAFGEVYNDKLIPTTISACKFLSEFNSIKDYITERLYNAFLAGSVVPLGSPREN
YENYIPADSFTHVEDYNSPELAKYLKEVDKNNKLIYSYFMWRKDFVNLPRFWESHA
CLACDHYKRQEKYSVGNLEKFMWN"
2192
/gene="FUT9"
/evidence=experimental
polyA_site 2235
/gene="FUT9"
/evidence=experimental
polyA_site 2310
/gene="FUT9"
/evidence=experimental
polyA_site 2466
/gene="FUT9"
/evidence=experimental
BASE COUNT 818 a 436 c 422 g 825 t
ORIGIN

Query Match 55.5%; Score 1205; DB 9; Length 2501;
Best Local Similarity 78.6%; Pred. No. 1.3e-194;
Matches 168; Conservative 0; Mismatches 380; Indels 75; Gaps 16;

QY 21 TTTAGAATGTGATACGATCGATGATAGACACCGTGGAGTAGTTCACGACATCTC- 79
DB 9 TTTAGAATGTGATACGATCGATGATAGACACCGTGGAGTAGTTCACGACATCTC- 68

QY 80 -TGCCTTCATGCTATGTTCTCTACATGAAAAATATGACATCAATCAAGGCATT 138
DB 69 GAGCTTCCCATGATATGTTCTCTATATGAAAAATATGACATCAATCAAGGCATT 128

QY 139 CTTCGCCATTTCTAATGCTCTGATCATCTCGGCTGCTTACATGTCATGCTCAT 198
DB 129 CTTCGCCATTTCTAATGCTCTGATCATCTCGGCTGCTTACATGTCATGCTCAT 188

QY 199 TATATCAAGCCCAACAGACGCTGGCTCAGTCAATGAGCTCCAGCTCTGCTG 258
DB 189 TATATCAAGCCCAACAGACGCTGGCTCAGTCAATGAGCTCCAGCTCTGCTG 248

QY 259 AAAATGAAAAATTTCTTCCACAAAAATGATTTATTTACGAACTACATCTGGTT 318
DB 249 AAAATGAAAAATTTCTTCCACAAAAATGATTTATTTAATGAACTACTATCTGG 308

QY 319 TGGGTATGCGCATTTGGGAGACCTTTGACCTTACATCTGCGCAAGCAATGTTCAATATC 378
DB 309 TGGGTATGCGCATTTGGGAGACCTTTGACCTTACATCTGCGCAAGCAATGTTCAATATC 368

QY 379 CAAGGTGCGCATCTCAACACAGACCGCTCATTTGACAAACAAATCCCATGCGCTGAT 438
DB 369 CAAGGTGCGCATCTCAACACAGACCGCTCATTTGACAAACAAATCCCATGCGCTGAT 428

QY 439 CACATAGAGACATCAGCTGGGATGCTGACTACTTACTCTCAGAGGCGAGCCACCTTT 498
DB 429 CACATAGAGACATCAGCTGGGATGCTGACTACTTACTCTCAGAGGCGAGCCACCTTT 488

QY 499 CAGAAATGATTTGGATGATTTAGAGTACCCACTCACACCCCCCAAAAGATGGCAT 558
DB 489 CAGAAATGATTTGGATGATTTAGAGTACCCACTCACACCCCCCAAAAGATGGCAT 548

QY 559 GAACATGTTGTAACCTGACTCACTAATTCGCGGATGCTAGATTCGAATGCTTAT 618
DB 549 GAACATGTTGTAACCTGACTCACTAATTCGCGGATGCTAGATTCGAATGCTTAT 608

QY 619 GGCCTCTTACGCGGTGAGCAACAAATCCCTTGTGTTGAAGTCCAGCAAGCAAGATTTG 678
DB 609 GGCCTCTTACGCGGTGAGCAACAAATCCCTTGTGTTGAAGTCCAGCAAGCAAGATTTG 668

QY 679 GTGTGCTGGGTTGTGATGATCTGGAACCTGAGCATGCCAGGTCAGATTTACACGAG 738
DB 669 GTGTGCTGGGTTGTGATGATCTGGAACCTGAGCATGCCAGGTCAGATTTACACGAG 728

QY 739 CTCAGCAAGATTTGAATTCACACACTTATGGCCAGCACTTGGGAGATATGCTGAACAT 798
DB 729 CTCAGCAAGATTTGAATTCACACACTTATGGCCAGCACTTGGGAGATATGCTGAACAT 788

QY 799 AAAAATCTGATCCCAACATATCTACTGTAATTTATCTTCAATTTGAAATCAAT 858
DB 789 AAAAATCTGATCCCAACATATCTACTGTAATTTATCTTCAATTTGAAATCAAT 848

QY 859 CACAAAGATTTACATCAGCAAAAAAGCTCTACATGATTTTGGCTGGTCACTACCTGTT 918
DB 849 CACAAAGATTTACATCAGCAAAAAAGCTCTACATGATTTTGGCTGGTCACTACCTGTT 908

QY 919 GTCTGGGTCATCTAGGCAAAAACTATGGAATTTATTTCCAGCTGATTCATTCAT 978
DB 909 GTCTGGGTCATCTAGGCAAAAACTATGGAATTTATTTCCAGCTGATTCATTCAT 968

QY 979 GTGGAAGATTTTAACTCTCCAGTAGTGAACAAATATCTGAGAGATGCAAAAAAC 1038
DB 969 GTGGAAGATTTTAACTCTCCAGTAGTGAACAAATATCTGAGAGATGCAAAAAAC 1028

QY 1039 AATAAGTGTACCTTAGTACTTATTAAGTGAAGAAAGATTTTACTGTAAACCTTACACGG 1098
DB 1029 AATAAGTGTACCTTAGTACTTATTAAGTGAAGAAAGATTTACTGTAAACCTTACACGG 1088

QY 1099 TTTTGGGAATCAATGATGCTGCTGCAATCCGATTTGAAAAGGCAATCAAGAAATATAG 1158
DB 1089 TTTTGGGAATCAATGATGCTGCTGCAATCCGATTTGAAAAGGCAATCAAGAAATATAG 1148

us-09-744-748-3.rge

[illegible]

ACCESSION sequence.
VERSION AL512406
KEYWORDS HTG.
SOURCE AL512406.14 GI:13897154
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Mammaliota; Metzger; Chorodeta; Canaliata; Vertebrata; Euteleostomi
TITLE 1 (bases 1 to 80247)
JOURNAL Pates, K.
Submitted (27-Apr-2001) Sanger Centre, Hinxton, Cambridgeshire,
Chr1:15A, UK. E-mail enquiries: humanyes@sanger.ac.uk
COMMENT On 30th 30, 2001 this sequence version replaced gi:1336709.
Where differences are found these are compared from overlapping clones
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
only a small overlap as described above, as we submit sequences with
this sequence were finished as follows unless otherwise noted: all
chemistry or covered by high quality data (i.e., phred quality >= 30).
as an attempt was made to resolve all regions of sequencing problems, such
as plasmids and repeats, all regions are covered by at least
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the features are with the source databases: Em: EMBL; Sw: SwissProt; Tr: TrEMBL; Wp: WormPeP; Information on the WormPeP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/WormPeP
was generated from part of bacterial clone contigs of human
Chromosome 6, from part of bacterial clone contigs of human
Group. Further information can be found at
http://www.sanger.ac.uk/Map/Chr6
Rpl1-50439 is from the library Rpl1-77815 is at 80148 in this sequence.
http://www.chori.org/bacpac/home.htm
Vektor: pAC3.6
IMPCR: pAC3.6
Rpl1-50439 is from the library Rpl1-77815 is at 80148 in this sequence.
This sequence is not the entire insert of clone
sections only once, except for a 100 base overlap.
The true left end of clone Rpl1-77815 is at 80148 in this sequence.
Location/Qualifiers
1. 80247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="Rpl1-50439"
/clone="Rpl1-77815"
18. 1422
/note="12 repeat: matches 875. .2419 of consensus"
174. 1834
/note="MIR repeat: matches 166. .258 of consensus"
2950. 12977
/note="14 copies 2 mer ca 100% conserved"
3411. 3569
/note="LTR16C repeat: matches 93. .364 of consensus"
4372. 4704
/note="LTR1A1 repeat: matches 5. .365 of consensus"
/note="match: 7821. .8406"
8126. 8184
/note="GSS: Em: A050748"
8185. 8596
/note="MER38B repeat: matches 272. .332 of consensus"
8779. 8817
/note="MER12D repeat: matches 1. .513 of consensus"
8819. 8972
/note="MER12D repeat: matches 504. .542 of consensus"
/note="MER58A repeat: matches 6. .162 of consensus"


```
misc_feature complement(9443..9891)
/note="match: GSS: Em:B90633"
misc_feature complement(9452..9883)
/note="match: GSS: Em:AQ119695"
repeat_region 9782..9894
/note="MIR repeat: matches 92..205 of consensus"
repeat_region 10089..10212
/note="MIR repeat: matches 135..262 of consensus"
misc_feature complement(10093..10671)
/note="match: GSS: Em:AQ237061"
repeat_region 10223..10282
/note="MIR repeat: matches 203..262 of consensus"
misc_feature complement(10315..10631)
/note="match: GSS: Em:AQ200259"
misc_feature 10805..11186
/note="match: GSS: Em:AQ072253"
repeat_region 11657..11844
/note="L2 repeat: matches 2546..2750 of consensus"
misc_feature complement(11659..12084)
/note="match: GSS: Em:AQ518361"
misc_feature complement(11674..11987)
/note="match: GSS: Em:AQ094678"
misc_feature 12083..12588
/note="match: GSS: Em:AQ691235"
repeat_region 12596..12810
/note="MIR repeat: matches 23..248 of consensus"
repeat_region 12920..12986
/note="L1M45 repeat: matches 6228..6295 of consensus"
repeat_region 12987..13169
/note="L1M45 repeat: matches 5852..6043 of consensus"
repeat_region 13168..13280
/note="L1M48 repeat: matches 5710..5832 of consensus"
repeat_region 13851..17218
/note="L1P1 repeat: matches 2787..6155 of consensus"
repeat_region 18903..18953
/note="M1R2CB repeat: matches 439..501 of consensus"
repeat_region 18984..19433
/note="M1R2CB repeat: matches 1..460 of consensus"
misc_feature complement(19108..19690)
/note="match: GSS: Em:AQ563935"
repeat_region 20683..20801
/note="Alu/dJ/FRAM repeat: matches 169..288 of consensus"
repeat_region 20802..20862
/note="Alu/dJ/FRAM repeat: matches 3..66 of consensus"
repeat_region 20912..21033
/note="MIR repeat: matches 114..237 of consensus"
repeat_region 21047..21128
/note="41 copies 2 mer ta 63% conserved"
repeat_region 21676..21859
/note="L2 repeat: matches 2130..2308 of consensus"
repeat_region 21890..23612
/note="L1P repeat: matches 841..2570 of consensus"
repeat_region 23601..26609
/note="L1P45 repeat: matches 3107..6145 of consensus"
repeat_region 26620..26789
/note="MER3A repeat: matches 1..174 of consensus"
repeat_region 28521..28627
/note="MER39 repeat: matches 5..111 of consensus"
repeat_region 28656..28837
/note="MER39b repeat: matches 54..235 of consensus"
repeat_region 28838..29142
/note="AluSc repeat: matches 4..308 of consensus"
repeat_region 29143..29494
/note="MER39b repeat: matches 235..549 of consensus"
repeat_region 30529..31344
/note="L1P43 repeat: matches 5331..6146 of consensus"
repeat_region 32234..32340
/note="L2 repeat: matches 2627..2750 of consensus"
repeat_region 32399..32578
/note="MIR repeat: matches 12..203 of consensus"
repeat_region 32816..32979
/note="MIR repeat: matches 91..259 of consensus"
misc_feature complement(33006..33450)
```

```
misc_feature /note="match: GSS: Em:AQ261227"
complement(33058..33444)
/note="match: GSS: Em:AQ040453"
repeat_region 33312..33596
/note="L1M4 repeat: matches 3473..3773 of consensus"
repeat_region 34856..34915
/note="30 copies 2 mer aa 68% conserved"
misc_feature complement(34987..35547)
/note="match: GSS: Em:AQ321598
match: STS: Em:G55571"
misc_feature complement(36488..36936)
/note="match: GSS: Em:B41073"
complement(36654..37090)
/note="match: GSS: Em:AQ772584"
misc_feature 37102..37591
/note="match: GSS: Em:AQ489567"
complement(37581..38221)
/note="match: GSS: Em:AQ18233"
repeat_region 37720..37772
/note="L2 repeat: matches 2657..2710 of consensus"
misc_feature complement(37778..38215)
/note="match: GSS: Em:AQ819773"
misc_feature 38218..38732
/note="match: GSS: Em:AQ599868"
repeat_region 38861..39164
/note="AluSc repeat: matches 1..305 of consensus"
repeat_region 39612..39653
/note="21 copies 2 mer ac 90% conserved"
repeat_region 40854..40909
/note="L2 repeat: matches 2655..2710 of consensus"
repeat_region 42169..42270
/note="MER3A repeat: matches 1..111 of consensus"
repeat_region 42411..42548
/note="MER3B repeat: matches 21..181 of consensus"
repeat_region 42780..42938
/note="MER3B repeat: matches 17..177 of consensus"
repeat_region 43229..43326
/note="MIR repeat: matches 40..148 of consensus"
repeat_region 4383..45428
/note="MER3B repeat: matches 18..62 of consensus"
misc_feature complement(47284..47921)
/note="match: STS: Em:G51995
match: GSS: Em:AQ269059"
misc_feature complement(47537..48011)
/note="match: STS: Em:G36899"
complement(47723..48014)
/note="match: STS: Em:G30049"
misc_feature 47972..48451
/note="match: GSS: Em:AQ336527"
complement(48518..48985)
/note="match: GSS: Em:AQ801834"
repeat_region 48597..48803
```

Query Match 53.6%, Score 1162.4; DB 9; Length 80247;
Best Local Similarity 78.7%; Pred. No. 1.4e-187;
Matches 1602; Conservative 0; Mismatches 361; Indels 73; Gaps 15;

```
Oy 107 GAAAAATTATGACATCAACATCCAAAGCATTCCTGCCCATTTCTATTCGTGCATCA 166
|||||
Db 35551 GAAAAATTATGACATCAACATCCAAAGAAATTCCTGCCCATTTTAAATGTCGATTA 35610

Oy 167 TCTTGCGCTGCTCATGGCATCTGCTCATTTATATTAAGCCACCAACAGCTGGTCT 226
|||||
Db 35611 TCTTGCGCTGTTTCAATGGCAGATGCTCTTCATTTATCATCAACACTACCAAGCTGATCT 35670

Oy 227 TCAGTCCAATGAGCTGCGAAGTCTGTGCTGCAAAATGCAAAATTTCTTCTCCACAAAAA 286
|||||
Db 35671 TCAGTCCAATGAGAAACACCCAGCTCTGTGCTGAAATGCAAAATCTTTTCCACAAAAA 35730

Oy 287 CTGATTATTTTAAGCAAACTCCATTCGTGGTTGGGTATGGCCATTGCGAGACCTTG 346
|||||
Db 35731 CTGATTATTTTAAGCAAACTACTATTCGTGTGGTGTGGCCATTGCGAGACCTTGG 35790
```



```

FEATURES
  source      Location/Qualifiers
              1. .1128
              /organism="Rattus norvegicus"
              /strain="Sprague-Dawley"
              /db_xref="taxon:10116"
              1. .1128
              /gene="Fut9"
              /gene="Fut9"
              /gene="Fut9"
              /codon_start=1
              /product="alpha1,3-fucosyltransferase IX"
              /protein_id="AAK16591.1"
              /db_xref="GI:13242184"
              /translacion="MTSTKGIILRPLVCIILGCPVACILYIKPTNSWVSPMESA
              SSVLKMKNEFSKRTDFENETTLILVWVPGQFEDLSCQAMENIQCHELTIDRLSYNK
              SHAVLHHHDIIMDLNLPQAPRPFQKWTMMLESPTEPKSGIEHLFTNLTLYR
              DSDIOPVGFPLTSTSPVEVPSKELICWVYVSNMNPBARVYNEHLSKSTIEHTY
              GQAFGEYVNDKNIIPRTSKKPLSYFENSIHNDYITEKIKYNAFLASVYVGLPSREN
              YENYIPADSPFIHVEDNPSELMKILKEYDKNNKLYLSTFNMKDFVNLPRWESHA
              CLACDHVRKHQEKSVGNLEKWFVN"
BASE COUNT      324 a      256 c      228 g      320 t
ORIGIN
Query Match      48.4%; Score 1049.6; DB 10; Length 1128;
Best Local Similarity 95.7%; Pred. No. 2.9e-168;
Matches 1079; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
67 CAGCAGCATCTGCTGCTTCATGCTATGTTCTCTACACTGAAAAATTTATGACATCAACA 126
1 CACAGCAGCATCTGCTTCATGATATGTTCTCTACCGGAAAAATTTATGACATCAACA 60
127 TCCAAAGCATCTGCGCCATTTCTAATGCTGATCATCTCGGCTGCTTCATGCA 186
61 TCCAAAGCATCTGCGCCATTTCTAATGCTGATCATCTCGGCTGCTTCATGCA 120
187 TGTCTGCTCATTTATATCAAGCCACCAAGAGTGGTCTTCAGTCCATGAGAGTGTGCA 246
121 TGTCTACTTATCTACATCAAGCCACCAAGAGTGGTCTTCAGTCCATGAGAGTGTGCA 180
247 AGTCTGCTGCTGAAAAATTTCTCTCCCAAAAAATTTATTTAAGAAACT 306
181 AGCTCTGCTGAAAAATTTCTCTCCAGGAAAAATTTATTTAATGAACCT 240
307 ACCATCTGCTGTTGGTATGCGCATTTGGGACAGACCTTTGACCTTACATCTGCCAAGCA 366
241 ACTATCTGCTGTTGGTATGCGCATTTGGGACAGACCTTTGACCTTACATCTGCCAAGCC 300
367 ATGTTCATATTCAGAGGTCGCATCTCACACAGACCGCTCATCTGTACAAACAAATCCCAT 426
301 ATGTTCATATTCAGAGGTCGCATCTCACACAGACCGCTCATCTGTGTACAAACAAATCCCAT 360
427 GCGGCTCTGATTCACCATAGAGACATCAGCTGGATCTGACTTAATCTTACCTGACAGGCC 486
361 GCGGCTCTGATTCACCATAGAGACATCAGCTGGATCTGACTTAATCTTACCTGACAGGCC 420
487 AGGCCACCCCTTTCAGAAATGATTTGATGAATTTAAGTACCCACCTCACACCCCCCA 546
421 AGGCCACCCCTTTCAGAAATGATTTGATGAATTTAAGTACCCACCTCACACCCCCCA 480
547 AAGAGTGGCATTCAGACATCTGTCAACCTGACTTAATGCGCGTATTTAGATATC 606
481 AAGAGTGGCATTCAGACATCTGTCAACCTGACTTAATGCGCGTATTTAGATATC 540
607 CAAGTCCCTTATGCTTCTTGAAGTGGAGACACAAATCCCTTGTGTTTGAAGTCCCAAG 666
541 CAAGTCCCTTATGCTTCTTGAAGTGGAGACACAAATCCCTTGTGTTTGAAGTCCCAAG 600
667 AAGGAGCAAGTGTGTGTGGTGTGTGAGTAACTGAGACCCCTGAGCATGCCAGGTCGCAAG 726
601 AAGGAGCAAGTGTGTGTGGTGTGTGAGTAACTGAGACCCCTGAGCATGCCAGGTCGCAAG 660
727 TATTAAAGAGCTCAGCAAGAGTATTGAATTCACACCTATGCGCAAGCATTTGGAGAA 786

```

```

Db 661 TATTAAAGAGCTCAGCAAGAGTATTGAATTCACACCTATGCGCAAGCATTTGGAGAA 720
Oy 787 TACGTGAACATATAAAATCGATTCACCACATATCTGTAATTTATTTATCTTATCTT 846
Db 721 TATGTGAATATATAAAATCGATTCACCACATATCTGTAATTTATTTATCTTATCTT 780
Oy 847 GAAAACTCAATTCAGAAAGATTCATACACAGAAAAAGCTCTACAAATGATTTTGGCTGT 906
Db 781 GAAAACTCAATTCAGAAAGATTCATACACAGAAAAAGCTCTACAAATGATTTTGGCTGT 840
Oy 907 TCAGTACTCTTGTGCTGGGTCATCTAGAGGAAAAATCATAGATTTATTTCCAGCTGAT 966
Db 841 TCGTACTCTTGTGCTGGGTCATCTAGAGGAAAAATCATAGATTTATTTCCAGCTGAT 900
Oy 967 TCATTCATTCATGAGAAAGATTTTAACTCCACAGTAGTATAGCAAAATTCGAAAGAA 1026
Db 901 TCATTCATTCATGAGAAAGATTTTAACTCCACAGTAGTATAGCAAAATTCGAAAGAG 960
Oy 1027 GTTGACAAAAAATAAGTTGTACCTTATGTTACTTTAACTGAGAAAGATTTTACTGTA 1086
Db 961 GTTGACAAAAAATAAGTTGTATATCTTATGTTACTTCACTGAGAGAGATTTTACTGTA 1020
Oy 1087 AACCTACACAGCTTTGGGAAATCAGATTCATGCTGCGCATGCGCATGTAAGAAAGCAT 1146
Db 1021 AACCTCCACAGCTTTGGGAAATCAGATTCATGCTGCGCATGCGCATGTAAGAAAGCAT 1080
Oy 1147 CAAGAATATATAGTCTGTGTGTAATTTAGAGAAATGCTTTGGAATTA 1194
Db 1081 CAAGAATATATAGTCTGTGTGTAATTTAGAGAAATGCTTTGGAATTA 1128

```

```

RESULT 8
AC100674/c
LOCUS      71396 bp      DNA      linear      HNG 22-NOV-2001
DEFINITION Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC100674
VERSION    AC100674.1 GI:17048040
KEYWORDS   HTG; HTGS; PHASE0.
SOURCE     house mouse.
ORGANISM   Mus musculus

```

```

REFERENCE
  AUTHORS   Birren, B., Linton, L., Nusbaum, C. and Lander, E.
  TITLE     Unpublished
  JOURNAL   2 (bases 1 to 71396)
  REFERENCE
    Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
    Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhallier, B.,
    Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
    Choquel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
    Cooke, P., Deavell, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
    Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Garlyna, S.,
    Glend, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
    Hagos, B., Heald, A., Horton, L., Hulme, W., Illey, I., Johnson, R.,
    Jones, C., Kamat, A., Karatas, A., Kellis, C., LaRoque, K.,
    Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
    Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
    McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
    Menais, J., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
    Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
    Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V.,
    Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
    Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
    Seaman, S., Severy, P., Spencer, B., Strange-Rhmann, N., Stojanovic, N.,
    Strauss, N., Sudramanian, A., Talamas, D., Testa, S., Theodore, J.,
    Topham, K., Travers, M., Travis, N., Triggillo, J., Vasilev, H.,
    Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
    Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

```

```

TITLE     Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
JOURNAL    Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT    All repeats were identified using RepeatMasker:

```

Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15932.
Center clone name: 167_K_24

* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 789 888: gap of 100 bp
* 889 1702: contig of 814 bp in length
* 1703 1802: gap of 100 bp
* 1803 2642: contig of 840 bp in length
* 2643 2742: gap of 100 bp
* 2743 3637: contig of 895 bp in length
* 3638 3737: gap of 100 bp
* 3738 4635: contig of 898 bp in length
* 4636 4735: gap of 100 bp
* 4736 5657: contig of 922 bp in length
* 5658 5757: gap of 100 bp
* 5758 6692: contig of 935 bp in length
* 6693 6792: gap of 100 bp
* 6793 7703: contig of 911 bp in length
* 7704 7803: gap of 100 bp
* 7804 8742: contig of 939 bp in length
* 8743 8842: gap of 100 bp
* 8843 9787: contig of 945 bp in length
* 9788 9887: gap of 100 bp
* 9888 10819: contig of 937 bp in length
* 10820 10919: gap of 100 bp
* 10920 11808: contig of 889 bp in length
* 11809 11908: gap of 100 bp
* 11909 12830: contig of 922 bp in length
* 12831 12930: gap of 100 bp
* 12931 13847: contig of 917 bp in length
* 13848 13947: gap of 100 bp
* 13948 14886: contig of 939 bp in length
* 14887 14986: gap of 100 bp
* 14987 15911: contig of 925 bp in length
* 15912 16011: gap of 100 bp
* 16012 16803: contig of 892 bp in length
* 16804 17003: gap of 100 bp
* 17004 17906: contig of 903 bp in length
* 17907 18006: gap of 100 bp
* 18007 18931: contig of 925 bp in length
* 18932 19031: gap of 100 bp
* 19032 19945: contig of 914 bp in length
* 19946 20045: gap of 100 bp
* 20046 20951: contig of 906 bp in length
* 20952 21051: gap of 100 bp
* 21052 21949: contig of 898 bp in length
* 21950 22049: gap of 100 bp
* 22050 22950: contig of 901 bp in length
* 22951 23050: gap of 100 bp
* 23051 23943: contig of 895 bp in length
* 23946 24045: gap of 100 bp
* 24046 24953: contig of 910 bp in length
* 24956 25055: gap of 100 bp
* 25056 25980: contig of 925 bp in length

* 25981 26080: gap of 100 bp
* 26081 27052: contig of 972 bp in length
* 27053 27152: gap of 100 bp
* 27153 28067: contig of 915 bp in length
* 28068 28167: gap of 100 bp
* 28168 29076: contig of 909 bp in length
* 29077 29176: gap of 100 bp
* 29177 30093: contig of 917 bp in length
* 30094 30193: gap of 100 bp
* 30194 31093: contig of 906 bp in length
* 31100 31199: gap of 100 bp
* 31200 32097: contig of 898 bp in length
* 32098 32197: gap of 100 bp
* 32198 33113: contig of 916 bp in length
* 33114 33213: gap of 100 bp
* 33214 34125: contig of 912 bp in length
* 34126 34225: gap of 100 bp
* 34226 35133: contig of 908 bp in length
* 35134 35233: gap of 100 bp
* 35234 36150: contig of 917 bp in length
* 36151 36250: gap of 100 bp
* 36251 37171: contig of 921 bp in length
* 37172 37271: gap of 100 bp
* 37272 38159: contig of 888 bp in length
* 38160 38259: gap of 100 bp
* 38260 39156: contig of 897 bp in length
* 39157 39256: gap of 100 bp
* 39257 40188: contig of 932 bp in length
* 40189 40288: gap of 100 bp
* 40289 41210: contig of 922 bp in length
* 41211 41310: gap of 100 bp
* 41311 42222: contig of 912 bp in length
* 42223 42322: gap of 100 bp
* 42323 43218: contig of 896 bp in length
* 43219 43318: gap of 100 bp
* 43319 44233: contig of 915 bp in length
* 44234 44333: gap of 100 bp
* 44334 45246: contig of 913 bp in length
* 45247 45346: gap of 100 bp
* 45347 46261: contig of 915 bp in length
* 46262 46361: gap of 100 bp
* 46362 47279: contig of 918 bp in length
* 47280 47379: contig of 100 bp
* 47380 48276: contig of 897 bp in length
* 48277 48376: gap of 100 bp
* 48377 49282: contig of 906 bp in length
* 49283 49382: gap of 100 bp
* 49383 50305: contig of 923 bp in length
* 50306 50405: gap of 100 bp
* 50406 51144: contig of 909 bp in length
* 51315 51414: gap of 100 bp
* 51415 52335: contig of 921 bp in length
* 52336 52435: gap of 100 bp
* 52435 53163: contig of 728 bp in length
* 53164 53263: gap of 100 bp
* 53264 54210: contig of 947 bp in length
* 54211 54310: gap of 100 bp
* 54311 55199: contig of 889 bp in length
* 55200 55299: gap of 100 bp
* 55300 56212: contig of 913 bp in length
* 56213 56312: gap of 100 bp
* 56313 57287: contig of 975 bp in length
* 57288 57387: gap of 100 bp
* 57388 58245: contig of 858 bp in length
* 58246 58345: gap of 100 bp
* 58346 59236: contig of 891 bp in length
* 59237 59336: gap of 100 bp
* 59337 60272: contig of 936 bp in length
* 60273 60372: gap of 100 bp
* 60373 61289: contig of 917 bp in length
* 61290 61389: gap of 100 bp
* 61390 62299: contig of 910 bp in length
* 62300 62399: gap of 100 bp

```

* 62400 63322: contig of 923 bp in length
* 63323 63422: gap of 100 bp
* 64323 64336: contig of 914 bp in length
* 64337 64436: gap of 100 bp
* 64437 65354: contig of 918 bp in length
* 65355 65454: gap of 100 bp
* 65455 66352: contig of 898 bp in length
* 66353 66452: gap of 100 bp
* 66453 67362: contig of 910 bp in length
* 67363 67462: gap of 100 bp
* 67463 68370: contig of 908 bp in length
* 68371 68470: gap of 100 bp
* 68471 69379: contig of 909 bp in length
* 69380 69479: gap of 100 bp

```

Query Match 40.7% Score 883; DB 2; Length 71396;

Best Local Similarity 88.0% Pred. No. 2,7e-140; Indels 1; Gaps 1;

```

Matches 906; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

QY 922 CTGGGTCATCTAGGAAAACTATGAAATTAATTCAGCTGATTCATTCATTCATG 981
DB 30210 CAGTGCAGACATCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 30151
QY 982 GAAGATTTTAACCTCCAGTGAGTACGAAATATCTGAGGAAGTTGACAAAAACAAT 1041
DB 30150 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAT 30091
QY 1042 AAGTTGACCTTGTACTTAACTGAGAAAGATTTTACTTAACCTACAC-GGTT 1100
DB 30090 AAGTTGACCTTGTACTTAACTGAGAAAGATTTTACTTAACTTAACTTAACTTAACT 30031
QY 1101 TTGGGAATCATGATGCTGCGATGACGACATGTAATAAAGGCATCAAGAAATATAGTC 1160
DB 30030 TTGGGAATCATGATGCTGCGATGACGACATGTAATAAAGGCATCAAGAAATATAGTC 29971
QY 1161 TGTGTGTAATTTAGAAATGCTTTGGAATTAAGTTCATTCATTCATTCATTCATTCAT 1220
DB 29970 TGTGTGTAATTTAGAAATGCTTTGGAATTAAGTTCATTCATTCATTCATTCATTCAT 29911
QY 1221 AGATTATGAGACATATATCAAGTTTTCAGATTAAGAAACAGCTGCATTTGGGGA 1280
DB 29910 AGATTATGAGACATATATCAAGTTTTCAGATTAAGAAACAGCTGCATTTGGGGA 29851
QY 1281 CCGTTAAATTTCTGCGCTCCCTGAGAGACATGTAATTTTGGTAAATTTTAAAG 1340
DB 29850 CCGTTAAATTTCTGCGCTCCCTGAGAGACATGTAATTTTGGTAAATTTTAAAG 29791
QY 1341 ATCAGAAATTTAGCAATCAGCAATTTGTTTAAATTTATCTGATATATGATTAAGAG 1400
DB 29790 ATCAGAAATTTAGCAATCAGCAATTTGTTTAAATTTATCTGATATATGATTAAGAG 29731
QY 1401 CACTGAAATTTATTTATTTGTACCTCATTTATTAACATTTGTTTACATTTTATAGTT 1460
DB 29730 CACTGAAATTTATTTATTTGTACCTCATTTATTAACATTTGTTTACATTTTATAGTT 29671
QY 1461 GACTGAAATTTATTTATTTGTACCTCATTTATTAACATTTGTTTACATTTTATAGTT 1520
DB 29670 GACTGAAATTTATTTATTTGTACCTCATTTATTAACATTTGTTTACATTTTATAGTT 29611
QY 1521 GGGAGTAAATTTATCTATCTTAATTAATGAGGAGCTTTTGGCAAGTATTAAGATAT 1580
DB 29610 GGGAGTAAATTTATCTATCTTAATTAATGAGGAGCTTTTGGCAAGTATTAAGATAT 29551
QY 1581 AGATTCTACTTGTATTAATGCTAACAAGAAATTAATTAATTAATTAATTAATTAAT 1640
DB 29550 AGATTCTACTTGTATTAATGCTAACAAGAAATTAATTAATTAATTAATTAATTAAT 29491
QY 1641 CTTTGACTTCAAGTGAACACTACTGACACTGATAGATTTTACTTGTACCAATTT 1700
DB 29490 CTTTGACTTCAAGTGAACACTACTGACACTGATAGATTTTACTTGTACCAATTT 29431
QY 1701 GTGAAGGAGCTAATTTGATGACAAATTAAGAAATTTGAAGCAAGAAATGTTATTAGA 1760
DB 29431 GTGAAGGAGCTAATTTGATGACAAATTAAGAAATTTGAAGCAAGAAATGTTATTAGA 1760

```

```

DB 29430 GTGAAGGAGCTAATTTGATGACAAATTAAGAAATTTGAAGCAAGAAATGTTATTAGA 29371
QY 1761 CTTACATTCATTTTATTAAGTGTGTTAGCAATGATGATCCCAATATTTTAACTTAA 1820
DB 29370 CTTACATTCATTTTATTAAGTGTGTTAGCAATGATGATCCCAATATTTTAACTTAA 29311
QY 1821 AGAAAAACAAGATTTTTCAGATTAACAATTTATATGCTTTTAAAGAGATTTGATTTGAT 1880
DB 29310 AGAAAAACAAGATTTTTCAGATTAACAATTTATATGCTTTTAAAGAGATTTGATTTGAT 29251
QY 1881 TTTGTGGGTTTACTCCCTTAATTAATTAAGATTAAGATTTTGTATTCCTTTGCAATTC 1940
DB 29250 TTTGTGGGTTTACTCCCTTAATTAATTAAGATTAAGATTTTGTATTCCTTTGCAATTC 29191
QY 1941 ATGAAGATA 1949
DB 29190 AGGGCGAGA 29182

```

RESULT 9
AB035906
LOCUS AB035906 1707 bp DNA linear VRT 05-JAN-2002
DEFINITION Gallus gallus gene for cfu9, complete cds.
ACCESSION AB035906
VERSION AB035906.1 GI:18146865
KEYWORDS cfu9.
SOURCE Gallus gallus brain DNA.
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE
1 (sites)
Kaneko, M., Nishihara, S., Kitano, T., Narimatsu, H. and Saitou, N.
The evolutionary history of glycosyltransferase genes
Unpublished
2 (bases 1 to 1707)
Kaneko, M., Saitou, N. and Kitano, T.
Direct Submission
Submitted (17-DEC-1999) Mika Kaneko, National Institute of
Genetics, Laboratory of Evolutionary Genetics, Yata 1111, Mishima,
Shizuoka 411-8540, Japan (E-mail: mkaneko@med.id.yamagata-u.ac.jp,
Tel: 81-559-81-6790, Fax: 81-559-81-6789)
Sequence updated (29-Feb-2000).

COMMENT
FEATURES
location/Qualifiers
1..1707
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="brain"
268..1347
/gene="cfu9-T1X"
268..1347
/gene="cfu9-T1X"
/codon_start=1
/product="cfu9"
/protein_id="BAB82489.1"
/db_xref="GI:18146866"
/translation="MTSRSKGIIRPFIVFIVIGGFMAFIIYIKPSWSMTSNVES
SSYLKMKSPFSKDNREYETIIVLMVRFPGTFLDTSQMPFNHIGLTLTDRSLYRK
SHAVLIHRDINMDLTNLPQARPEPQKIMWNLESPTHTOKSGIEHLFNLTYLR
DSDIQVPYGSIMIVGATVEPVSKEINVCWVSWNDEHARVRYNLSYIEITHY
GOAFEDYVNDKNIIPITSSCKEYFSEMSIKDYITEKLYNALLAGSVPIVIGSREN
YENTIPADSFTHVEDFLSPRELAETYLMLDKNNKKYLSFYFMRBDSVHLPKPESH
CLACDHVRHDEYSIGLEKFEFNN"

gene
CDS

BASE COUNT 517 a 338 c 329 g 521 t 2 others
ORIGIN

Query Match 34.1% Score 740; DB 5; Length 1707;
Best Local Similarity 72.2% Pred. No. 6.9e-116;
Matches 1002; Conservative 1; Mismatches 356; Indels 29; Gaps 2;

```

QY 108 AAAAAATTAATGACATCAACATCCAAAGGCAATTTGCGCCATTTCTTAATCTGATCAT 167
DB 261 ACAAATTAATGACATCAACATCTAAAGGAATTTTCGGGCATTTTAAATGTCCTCATGCT 320

```

RESULT	DOCU	ACCEN	VERSION	KEYWORD	SOURCE	ORGANISM
1242	AGTTTTCGATTAAGAAACACCGCTGCA					
1401	AGTGAGAGTAAGGAATTAAGCAAAAGAAAGAGAGATGATACCTGACAGAACCTTACG					
1279	CACCGTTTAAATTTTCTGCGCCCTCGTTGAGGACCATGATATATTTGGGAAATTTTAAAG					
1461	CAAGGTATATTCAGCCATCTCTCTTGGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1339	AGATGAGATTAAGCAATCACTCATTTGGTGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1521	ATTATGAGAGAGATCAATCACTCATTTGGTGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1399	AGCAGCGAATTAATTTATTTTCTGCTCATTTTAAACATGTTTATTCATTTTAAAG					
1581	GCATCGAAGAAATATATGATTTTCTGCTCATTTTAAACATGTTTATTCATTTTAAAG					
1459	TTGAGTCT 1466					
1641	TGAGATAT 1648					
1242	AGTTTTCGATTAAGAAACACCGCTGCA					
1401	AGTGAGAGTAAGGAATTAAGCAAAAGAAAGAGAGATGATACCTGACAGAACCTTACG					
1279	CACCGTTTAAATTTTCTGCGCCCTCGTTGAGGACCATGATATATTTGGGAAATTTTAAAG					
1461	CAAGGTATATTCAGCCATCTCTCTTGGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1339	AGATGAGATTAAGCAATCACTCATTTGGTGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1521	ATTATGAGAGAGATCAATCACTCATTTGGTGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1399	AGCAGCGAATTAATTTATTTTCTGCTCATTTTAAACATGTTTATTCATTTTAAAG					
1581	GCATCGAAGAAATATATGATTTTCTGCTCATTTTAAACATGTTTATTCATTTTAAAG					
1459	TTGAGTCT 1466					
1641	TGAGATAT 1648					
1242	AGTTTTCGATTAAGAAACACCGCTGCA					
1401	AGTGAGAGTAAGGAATTAAGCAAAAGAAAGAGAGATGATACCTGACAGAACCTTACG					
1279	CACCGTTTAAATTTTCTGCGCCCTCGTTGAGGACCATGATATATTTGGGAAATTTTAAAG					
1461	CAAGGTATATTCAGCCATCTCTCTTGGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1339	AGATGAGATTAAGCAATCACTCATTTGGTGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1521	ATTATGAGAGAGATCAATCACTCATTTGGTGCTATTCATTTATTTTATTTTAAAGATTTTAA					
1399	AGCAGCGAATTAATTTATTTTCTGCTCATTTTAAACATGTTTATTCATTTTAAAG					
1581	GCATCGAAGAAATATATGATTTTCTGCTCATTTTAAACATGTTTATTCATTTTAAAG					
1459	TTGAGTCT 1466					
1641	TGAGATAT 1648					

* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 788: contig of 788 bp in length
789 888: gap of 100 bp
889 1702: contig of 814 bp in length
1703 1802: gap of 100 bp
1803 2642: contig of 840 bp in length
2643 2742: gap of 100 bp
2743 3637: contig of 895 bp in length
3638 3737: gap of 100 bp
3738 4635: contig of 898 bp in length
4636 4735: gap of 100 bp
4736 5657: contig of 922 bp in length
5658 5757: gap of 100 bp
5758 6692: contig of 935 bp in length
6693 6792: gap of 100 bp
6793 7703: contig of 911 bp in length
7704 7803: gap of 100 bp
7804 8742: contig of 939 bp in length
8743 8842: gap of 100 bp
8843 9787: contig of 945 bp in length
9788 9887: gap of 100 bp
9888 10819: contig of 932 bp in length
10820 10919: gap of 100 bp
10920 11808: contig of 889 bp in length
11809 11908: gap of 100 bp
11909 12830: contig of 922 bp in length
12831 12930: gap of 100 bp
12931 13847: contig of 917 bp in length
13848 13947: gap of 100 bp
13948 14886: contig of 939 bp in length
14887 14986: gap of 100 bp
14987 15911: contig of 925 bp in length
15912 16011: gap of 100 bp
16012 16903: contig of 892 bp in length
16904 17003: gap of 100 bp
17004 17906: contig of 903 bp in length
17907 18006: gap of 100 bp
18007 18931: contig of 925 bp in length
18932 19031: gap of 100 bp
19032 19945: contig of 914 bp in length
19946 20045: gap of 100 bp
20046 20951: contig of 906 bp in length
20952 21051: gap of 100 bp
21052 21949: contig of 858 bp in length
21950 22049: gap of 100 bp
22050 22950: contig of 901 bp in length
22951 23050: gap of 100 bp
23051 23945: contig of 895 bp in length
23946 24045: gap of 100 bp
24046 24955: contig of 910 bp in length
24956 25055: gap of 100 bp
25056 25980: contig of 925 bp in length
25981 26080: gap of 100 bp
26081 27052: contig of 972 bp in length
27053 27152: gap of 100 bp
27153 28067: contig of 915 bp in length
28068 28167: gap of 100 bp
28168 29076: contig of 909 bp in length
29077 29176: gap of 100 bp
29177 30093: contig of 917 bp in length
30094 30193: gap of 100 bp
30194 31099: contig of 906 bp in length

31100 31199: gap of 100 bp
31200 32097: contig of 898 bp in length
32098 32197: gap of 100 bp
32198 33113: contig of 916 bp in length
33114 33213: gap of 100 bp
33214 34125: contig of 912 bp in length
34126 34225: gap of 100 bp
34226 35133: contig of 908 bp in length
35134 35233: gap of 100 bp
35234 36150: contig of 917 bp in length
36151 36250: gap of 100 bp
36251 37171: contig of 921 bp in length
37172 37271: gap of 100 bp
37272 38159: contig of 888 bp in length
38160 38259: gap of 100 bp
38260 39156: contig of 897 bp in length
39157 39256: gap of 100 bp
39257 40188: contig of 932 bp in length
40189 40288: gap of 100 bp
40289 41210: contig of 922 bp in length
41211 41310: gap of 100 bp
41311 42222: contig of 912 bp in length
42223 42323: gap of 100 bp
42323 43218: contig of 866 bp in length
43219 43318: gap of 100 bp
43319 44233: contig of 915 bp in length
44234 44333: gap of 100 bp
44334 45246: contig of 913 bp in length
45247 45346: gap of 100 bp
45347 46261: contig of 915 bp in length
46262 46361: gap of 100 bp
46362 47279: contig of 918 bp in length
47280 47379: gap of 100 bp
47380 48276: contig of 897 bp in length
48277 48376: gap of 100 bp
48377 49282: contig of 906 bp in length
49283 49382: gap of 100 bp
49383 50305: contig of 923 bp in length
50306 50405: gap of 100 bp
50406 51314: contig of 909 bp in length
51315 51414: gap of 100 bp
51415 52335: contig of 921 bp in length
52336 52435: gap of 100 bp
52436 53163: contig of 728 bp in length
53164 53263: gap of 100 bp
53264 54210: contig of 947 bp in length
54211 54310: gap of 100 bp
54311 55199: contig of 889 bp in length
55200 55299: gap of 100 bp
55300 56212: contig of 913 bp in length
56213 56312: gap of 100 bp
56313 57287: contig of 975 bp in length
57288 57387: gap of 100 bp
57388 58245: contig of 858 bp in length
58246 58345: gap of 100 bp
58346 59236: contig of 891 bp in length
59237 59336: gap of 100 bp
59337 60272: contig of 936 bp in length
60273 60372: gap of 100 bp
60373 61289: contig of 917 bp in length
61290 61389: gap of 100 bp
61390 62299: contig of 910 bp in length
62300 62399: gap of 100 bp
62400 63322: contig of 923 bp in length
63323 63422: gap of 100 bp
63423 64336: contig of 914 bp in length
64337 64436: gap of 100 bp
64437 65354: contig of 918 bp in length
65355 65454: gap of 100 bp
65455 66352: contig of 898 bp in length
66353 66452: gap of 100 bp
66453 67362: contig of 910 bp in length
67363 67462: gap of 100 bp

us-09-744-748-3.rge

[illegible]

LOCUS	AF288369	1264 bp	DNA	linear	VRT 02-DEC-2003
DEFINITION	Gallus gallus alpha (1,3) fucosyltransferase gene, complete cds.				
ACCESSION	AF288369				
VERSION	AF288369.1	GI:17224489			
KEYWORDS					
SOURCE	chicken.				
ORGANISM	Gallus gallus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.				
AUTHORS	1 (bases 1 to 1264)				
TITLE	Foley, A.C., Schmitz, B., Stern, C.D. and Streit, A.				
JOURNAL	Expression of fucosyltransferases in chick embryos				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 1264)				
TITLE	Foley, A.C., Schmitz, B., Stern, C.D. and Streit, A.				
JOURNAL	Direct Submission				
FEATURES	Submitted (19-JUL-2000) Genetics and Development, Columbia University, 701 West 168th Street, New York, NY 10032, USA				
SOURCE	Location/Qualifiers				
CDS	1..1264				
	/organism="Gallus gallus"				
	/db_xref="taxon:9031"				
	<1..1062				
	/note="fuc"				
	/codon_start=1				
	/product="alpha (1,3) fucosyltransferase"				
	/protein_id="FAL36994.1"				
	/db_xref="GI:17224490"				
	/translation="SMKKILFLFLFSFISFSLRSLRPFPPEDYNSLSTERLRL AGSISMLPKSRMLTILMTWTFPGHRENSCELEFDPMCHFTTNMSWSDRAVLH HHRVYWDREKLSRIPLRISQRMWTFWENSSHPNIGAMDLNLTMSYRSDIFLTF PYGLQILDSREKSPISPPKTKVMAMVSNMRRSRVYKYEELKKHLPVYDGRHLE LPMQMLHQTFVQSNFYLAENSEKDYITTELMNLAASSGIIPVLGPPRENYHFLLE PDSITVHDDFASAGELAOYLQELSRDIEKRYQYRWRRWLKPVLSIGMAYVCACHH LQREASAHVYPNLSKMFV"				
BASE COUNT	286 a	355 c	319 g	304 t	
ORIGIN					
Query Match	10.2%;	Score 220.6;	DB 5;	Length 1264;	
Best Local Similarity	57.2%;	Pred. No. 5.8e-28;			
Matches 441;	Conservative 0;	Mismatches 324;	Indels 6;	Gaps 2;	
OY	307	ACCAATTCGTTGGTGGATATGCGCATTTGGGGCAGACCTTTGACCTTACATTCCTGCCAAGCA	366		
DB	175	ACCAATTCGTTGGTGGATATGCGCATTTGGGGCAGACCTTTGACCTTACATTCCTGCCAAGCA	234		
OY	367	ATGTTCAATATTCAGAGGTGCCATCTCACAAACAGACCGCTCATTTGTACAAACAATCCCAT	426		
DB	235	CTCTTTGACACCCCGCAATGCCCATCTTTCACACCAACGCGACGCTGCTCCCAAAAGCAGAT	294		
OY	427	GCGGTCGATTTCCACCATATAGACATCATGACGCGGATCTGCATTACTTACCTCAGAGGCC	486		
DB	295	GCTGATATTTATGACATCACACAGGATGTGATTTATGGGACAGACAGAGAGGCTGTCCCGATCC	354		
OY	487	AGGCCACCCCTTTCAGAAATGATTTTGGATTTAGATGATGACACCATCTACACCCCCCAA	546		
DB	355	AGGCCACCATCCACGCGCTGATCTGGTTCAACATGAGAGTCCCCCAAGTATCTTCCAAAC	414		
OY	547	AAGAGTGCATTCAGACCTTGTCAACCTGACTTAACCTTATCGCGGTGATTCAGATATC	606		
DB	415	TTAGTTCGCATGACACCACTCTTCAACCTGCATATGCTACCGGAGAGACATCGGACATC	474		
OY	607	CAAGGCGCTTATGCGCTTTTGACGGTGAACCAATCCCTTGTGTGTTGAAATGCCAAGC	666		
DB	475	TTTACACCCCTTATGCGGAGCTGACGCTGATGACGCGGACCCACGCTTACGATATCCACCC	534		
OY	667	AAGGAGAGTTGATGCTGCTGGTTGTGATTAATTCGAAACCTTGAGCATGCGAGGTCAG	726		
DB	535	AAGACCAAGATATGTCCTGGTGTGATTAATTTGGAAGACCAAGCTCCACCGGGTGAAG	594		
OY	727	TTATTCAACAGAGCTCAGCAAGATATTTGAAATTCACACACCTATGCGCAAGCATTTGGAGAA	786		

Db 595 TACTACGAGGAGCGTGAAGAAACACATCCCTCGTGCATGTGTACGGGGCGCAACAC---TTG 651

Qy 787 TACGTGAACGATAAAATCTGATTCCCAACCATATCTACTGTGTAATTTTATCTTCATT 846

Db 652 CCCCTGTCATGTGACATGCTCCACCAAAACCGTATCCAGTACAACTTCTACTTGACCTTT 711

Qy 847 GAAACCTCAATTCCAAAGATTACATGCACGAAAACCTCTAC--ATGCAATTTTGGCT 903

Db 712 GAGAACTAGAGACACAAAGACTACATCACTGGAAGAGCTCTGGAGAAATCCCTGTCCTT 771

Qy 904 GATTAGTACCTGTTGTCTCTGGTCCATCTAGGAAAACTATGAGAAATATATTCACAGT 963

Db 772 GGCACCATCCCAAGTGTCTCTGGGGCCACTCGAGAAAACATATGAGACACTTCTGGCCCT 831

Qy 964 GATTCAATTCATTCATGTGGAAGATTTTAATCTCCAGTAGTACGACAAATATCTGAAG 1023

Db 832 GACTCTTCATTCACAGTGCATCTTTGGCAGATGCTGGTGAACCTGGCGGAGTACCTCGAG 891

Qy 1024 GAAGTTGCAAAAAACAATATAGTTTACTTACTTACTTACTTACTGCGAAG 1074

Db 892 GAAGTACGACGAGACACGAGAAATGACGAGCTCACTTCAGTGGCGCAGG 942

RESULT 15

LOCUS	AB035907	1092 bp	DNA	linear	VRT 05-JAN-2002
DEFINITION	AB035907	Xenopus laevis gene for xFTI, complete cds.			
ACCESSION	AB035907				
VERSION	AB035907.1	GI:18146867			
KEYWORDS	xFTI.				
SOURCE	Xenopus laevis bone marrow DNA.				
ORGANISM	Xenopus laevis				
REFERENCE	1 (sites)				
AUTHORS	Kaneko,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.				
TITLE	The evolutionary history of glycosyltransferase genes				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1092)				
AUTHORS	Kaneko,M., Saitou,N. and Kitano,T.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-DEC-1999) Mka Kaneko, National Institute of Genetics, Laboratory of Evolutionary Genetics; Yata J111, Mishima, Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp, Tel:81-559-81-6790, Fax:81-559-81-6789)				
FEATURES	Location/Qualifiers				
Source	1..1092				
gene	/organism="Xenopus laevis"				
CDS	/db_xref="taxon:8355"				
	/tissue-type="bone marrow"				
	1..1092				
	/gene="xFTI"				
	1..1092				
	/gene="xFTI"				
	/codon_start=1				
	/product="xFTI"				
	/protein_id="BAB82490.1"				
	/db_xref="GI:18146868"				
	/translation="MRGDSLISVRLRQWRLCRSRRLSLRLLAASLTCTVAVVYVCSOGLAGINFLPLMGTTTPPKQVTVLIWPEPKRRRLDCHMEFNVGCHITTNNSLYVEDAIHMRHSDIDSEDLPPRRROSSOKWIMNESPESHSPWLASLGGIFNMWTDYSYRDLIDAKEDNIIVTVSEKTKYTLAFENSLHTDITTEKLMRAFSNALPIYMGPSRYNEMETIPNSFSLHVDSESPRLAVLAIKHLIDNNILHYRYETWKKRYDVHVSFWMEHCTCAQSVKAAQNRHTRASDLGWFES"				
BASE COUNT	256 a 281 c 294 g 261 t				
ORIGIN					

Query Match 9.5%; Score 206.6; DB 5; Length 1092;
 Best Local Similarity 56.1%; Pred. No. 1,4e-25;
 Matches 474; Conservative 0; Mismatches 359; Indels 12; Gaps 4;

Tue Oct 8 10:18:15 2002

us-09-744-748-3.rge

Page 20

QY 298 ACGAAACCTACCATTCGTGGTTGGGTATGCGCATTTGGGAGACCTTTGACCTTACATCC 357
 Db 199 AAGCAATCACTGTCTTAATATGATGATGAGACCATTTTGGCAAAAGCGAAGACTTGGGAT 258
 QY 358 TGGCAAGCAATGTTCAATATATCAAGGGGTGGCCATCTCCAGCAACAGACCGCTCATTTATAC 417
 Db 259 TGGCACTACTCTTATCAATATGATGAGGGGTGGCCATCAACACACAGAGAGCTTTACAG 318
 QY 418 AATCCCAATGGGGTCTGATTTACCAATAGACATCACTGGGTATGCTGACTACTTACCT 477
 Db 319 GAGGAGAGCGCATTTATGTTCCATCCACACCGGGACATT---TGTATTTCCATGACCTTCT 375
 QY 478 CAGCAGCGCAGCGCACCTTTGAGAAATGATTTGGATGAAATTTAAGACACCCACTAC 537
 Db 376 TTCCGAAGGCGCCAGAGCTCCCAAAAGTGGATGGATGGAACCTTGCATCTCCGTCAT 435
 QY 538 ACCCGCCAAAAGTGGCATGGAACACTGTTCAACCTGACTTAACCTTATGCGCGGAT 597
 Db 436 TCGCCTCGGCTGGAGTCTGGGGGATCTTCACTGAGACCATGTCCTCAAGGTGGAC 495
 QY 598 TCGATATCCAAAGTCCCTTAATGGCTCTTGACGGAGACACAAATGCCCTTTGGTTTGA 657
 Db 496 TCGGATATCTCAATGCTCCCTACGGCTCTT---ATTCTCAGAAACCTGCCAAATATGTC 612
 QY 658 GTGCCAACAAGAGGAGGATTTGGTGTGCTGGTTGGTGTGATATGCAAAACCTGAGCATGCC 717
 Db 553 CTGGCGCGCAGAGAAACCTGCTGGTGGCGGGGTGATGACAACTGCAAGCAGGACCCACGA 777
 QY 718 AGGTCAGATTTATCAACAGCTGGCCAGCAAGATTTGAAATTCACACCCATGAGCCCAAGA 837
 Db 613 AGGGTCAGATTTATCAACAGCTGAGCACTGAGCACTAGCTGGAATATGAGTTTAGGGC---GG 894
 QY 778 TTGGGGAATACGTGAACGATTAATAATTCATATCCCAACCATATGCTTGAATTTAT 954
 Db 670 TATGGCTGGATTTTAAAGGAGCAACATCTGACATGACCGCTGGAGTTAAATTTTAC 1014
 QY 838 CTTTATTTGAAACCTCAATTCACAAAGATTCATCAACAGAAAGCTCT---ACATGCA 1074
 Db 730 TTGGCTTTTGAATCTCTTGCCACAGGATTAACATCAACGGAAGAGCTGGAGGAACCG 1134
 QY 895 TTTTGGCTGCTTCACTAGCTCTTTGCTGGTGCATCTTAGGGAACCACTATGAGAAATAT 1193
 Db 790 TTATATTCAAACCCATCCCATCTGATCTGATGGGCCAGCGCATATATCTAGCAATATTC 1258
 QY 955 ATTCCAGCTGATTCATTCATTCATGAGATTTTAACTCTCCAGAGGATTTAGCAAA 1318
 Db 850 ATACCCCGCAGCTTTTCACTCACTGAGCAATTTCTCCAGCCCCAGAGAGCTGGCCAG 1378
 QY 1015 TATCTAAGAGATTTGACAAAACCAATTAAGTTGGTCTTACTTACTTAAGTGGAAAG 1438
 Db 910 TACCTGAGACACTGGATTAATAAAACATCTACTGTACCGCAGGATTTTCAAGTGAAG 1498
 QY 1075 GATTTTACTGTAAACCTACACGCGTTTGGAGATCAACATCACTGCGCATGCCAAT 1558
 Db 970 CGTTACACAGCTGACAGCTCTTCTTGGGACAGCATTAAGTACACAGCGTGGCAGGC 1618
 QY 1135 GTAAA 1139
 Db 1030 GTCAA 1034

Search completed: October 6, 2002, 04:09:47
Job time : 3282.24 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 243.519 Seconds
(without alignments)
15299.444 Million cell updates/sec

Title: US-09-744-748-3

Perfect score: 2170
Sequence: 1 cttccctcgtggtgcgc.....aaaaaaaaaaaaaaaa 2170

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2170	100.0	2170	21	AAZ92645 Murine alpha-1.3-f
2	2036	93.8	2036	21	AAZ82672 Murine alpha-1.3-f
3	1206.8	55.6	2676	21	AAZ82647 Human alpha-1.3-fu
4	1092.2	50.3	2822	21	AAZ92646 Human alpha-1.3-fu
5	129	5.9	1814	18	AAI59505 Murine myeloid-1in
6	128.6	5.9	3594	18	AAI85219 Mouse alpha-fucosyl
7	128.2	5.9	1086	15	AAO56912 Lewis blood group
8	128.2	5.9	2042	18	AAI80111 Fucosyltransferase
9	128.2	5.9	2043	12	AAQ13350 Glycosyltransferase

10	128.2	5.9	2043	15	AAO56906 DNA encoding a gly
11	128.2	5.9	2043	18	AAI76769 Human alpha 1.3/4
12	128.2	5.9	2043	18	AAI61675 Human alpha1.3/1,
13	126.6	5.8	1316	15	AAO56911 DNA encoding a gly
14	126.6	5.8	1316	18	AAI61679 Human alpha1.3)-f
15	124.6	5.7	1654	15	AAO56905 pCDNA1-alpha-(1-3)
16	124.6	5.7	1654	18	AAI61680 Human alpha(1.3)-f
17	123	5.7	1126	21	AAI21135 Human low adenosin
18	123	5.7	1126	21	AAI21133 Human adenosine re
19	123	5.7	6944	21	AAI21137 Human low adenosin
20	123	5.7	6944	21	AAI35015 Human adenosine re
21	122.8	5.7	1701	15	AAQ77732 Human alpha-1.3-fu
22	122.8	5.7	1701	21	AAI21136 Human low adenosin
23	122.8	5.7	1701	21	AAI35014 Human adenosine re
24	120	5.5	1256	21	AAI21134 Human low adenosin
25	120	5.5	1256	21	AAI35012 Human adenosine re
26	120	5.5	1400	13	AAO31436 Encodes a HeLa cel
27	120	5.5	1488	15	AAO36910 DNA encoding a gly
28	120	5.5	2175	12	AAQ14382 Human 7.2 encoding
29	120	5.5	2175	24	AAI37082 Human cDNA clone 7
30	120	5.5	2861	12	AAO14383 Human low adenosin
31	120	5.5	2861	21	AAI21133 Human adenosine re
32	120	5.5	2861	21	AAI35011 Human adenosine re
33	120	5.5	2861	24	AAI37083 Human cDNA clone 1
34	120	5.5	3647	12	AAO33333 GDP-Fuc:beta-D-Gal
35	120	5.5	3647	15	AAO56909 DNA encoding a gly
36	120	5.5	3647	18	AAI61678 Human alpha(1.3)-f
37	118.4	5.5	2175	11	AAO6691 CDX, a MILA involv
38	116.8	5.4	2134	18	AAI59506 Human myeloid deri
39	116	5.3	795	17	AAI3798 Human gene regulat
40	87	4.0	87	21	AAI2671 Human immune syste
41	80.4	3.7	19380	24	AAI561426 Human immune syste
42	78	3.6	15548	24	ABL34155 Human immune syste
43	77.4	3.6	15373	24	ABL32467 Human immune syste
44	74.6	3.4	9810	24	ABL32427 Human immune syste
45	74.2	3.4	7167	24	ABL32400 Human immune syste

ALIGNMENTS

RESULT 1	
AAZ92645	
ID	AAZ92645 standard; CDNA: 2170 BP.
XX	
AC	AAZ92645;
XX	
DT	05-JUN-2000 (first entry)
XX	
DE	Murine alpha-1.3-fucosyltransferase cDNA.
XX	
KW	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW	brin; kidney; recombinant expression; transgenic animal; knockout
KW	animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
KW	treatment; cancer; murine; mouse; ss.
XX	
OS	Mus sp.
XX	
EH	
FT	key
FT	Location/Qualifiers
FT	CDS 115..1194
FT	/*tag= a
XX	/product= "Murine alpha-1.3-fucosyltransferase"
PN	MO200006708-A1.
XX	
PD	10-FEB-2000.
XX	
PF	29-JUL-1999; 99WO-JP04092.
XX	
PR	29-JUL-1998; 98JP-0213823.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	

Tue Oct 8 10:18:16 2002

us-09-744-748-3.rng

Page 2

PI Narimatsu H, Kudo T, Sasaki K:
XX
DR WPI: 2000-183120/16.
DR P-PDB: AAY80935.
XX
PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer. -
XX
XZ Claim 4; Page 127-134; 172pp; Japanese.

The invention relates to a novel alpha-1,3-fucosyltransferase which transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglucosamine (generating the Lewis x or y epitope). It does not transfer a fucose moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglucosamine and therefore does not generate the sialyl-Lewis x epitope. The invention also relates to DNA sequences encoding alpha-1,3-fucosyltransferase and expression vectors and host cells comprising these DNA sequences. The invention additionally encompasses the preparation of alpha-1,3-fucosyltransferase via the culture of transformed cells or by expression of the protein in a transgenic animal; antibodies which recognize alpha-1,3-fucosyltransferase; methods for screening potential inhibitors or potentiators of alpha-1,3-fucosyltransferase activity or expression; the preparation of compounds having fucose-containing sugar chains by use of the protein; and knockout non-human animals lacking alpha-1,3-fucosyltransferase. Alpha-1,3-fucosyltransferase has a similar substrate range to the known FUC-IV and is expressed mainly in brain and kidney tissues. Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies, potentiators and inhibitors may be used for the treatment and diagnosis of diseases of the brain and kidney, and of cancers. They may be used for the identification of substances which affect the activity or expression of alpha-1,3-fucosyltransferase; such substances may be used therapeutically. The knockout animals can be used to study the mechanisms of action and expression of alpha-1,3-fucosyltransferase. Sequence AA92845 represents cDNA encoding murine alpha-1,3-fucosyltransferase (AA92845), and sequences AA92870-92873 are murine alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid amplification of cDNA ends).

Sequence 2170 BP; 715 A; 395 C; 368 G; 692 T; 0 other;

Query Match	100.0%	Score 2170;	DB 21;	Length 2170;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2170; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

OY	1	CTCTCCCTGCTGGGATGCGCTTTTGAATGTGATTAATCAGCATGGAATGACCGCTGGA	60
Db	1	CTTTCCTCCGCTGGGGATTCCTCTTGAATGTGATTAATCAGCATGGAATGACCGCTGGA	60
OY	61	GTAATTCAGCACTCATCTGCTGGCTTCATGATATTTCTACACAGTGAATAAATATATACA	120
Db	61	GTAGTTCAGCACTCATCTGCTGGCTTCATGATATTTCTACACAGTGAATAAATATATACA	120
OY	121	TCAACATCCAAAGGCAATCTTGTGGCAATTTCTTAATGCTGCAATCAATCTGGGTGCTTC	180
Db	121	TCAAAATCCAAAGGCAATTTTGTGCCATTTCTTAATGCTGCAATCAATCTGGGTGCTTC	180
OY	181	ATGCGATGTGCTCATTTTATATCAAGCCACCAACAGCTGGGTCTTCAGTCCAAATGAG	240
Db	181	ATGGGATGTGCTCATTTTATATCAAGCCACCAACAGCTGGGTCTTCAGTCCAAATGAG	240
OY	241	TCGCAATCTGTGCTGAGAAATGAAAATATTTCTCTGACAGAAAACGTATTAATTTTAAAC	300
Db	241	TCGCAATCTGTGCTGAGAAATGAAAATATTTCTCTGACAAAACGTATTAATTTTAAAC	300
OY	301	GAAATACCAATCTGCTGTTGGGTATGAGCCATTTTGGGAGAGACTTTTAACTTACATCTG	360
Db	301	GAAATACCAATCTGCTGTTGGGTATGAGCCATTTTGGGAGAGACTTTTAACTTACATCTG	360
OY	361	CAGCAATGTTCAATATCCAGAGGTGGCATCTTCACAGACGCGTCATGTGACACAAA	420
Db	361	CAGCAATGTTCAATATCCAGAGGTGGCATCTTCACAGACGCGTCATGTGACACAAA	420

Qy	421	TCACATCGGCTCTGATTTACATATAGACATCAGCTGGGATCTGACTAACTACTGAC	480
Db	421	TCACATCGGCTCTGATTTACATATAGACATCAGCTGGGATCTGACTAACTACTGAC	480
Qy	481	CAGGCCAGGCCACCTTTCAGAAATGGATTGGATGTAATTTAGAGTCACCCACTCACCC	540
Db	481	CAGGCCAGGCCACCTTTCAGAAATGGATTGGATGTAATTTAGAGTCACCCACTCACCC	540
Qy	541	CCCCAAAAAGATGGCATTTGAACACTTGTCAACCTGACCTCTAACTATATGCGCCGTATCA	600
Db	541	CCCCAAAAAGATGGCATTTGAACACTTGTCAACCTGACCTCTAACTATATGCGCCGTATCA	600
Qy	601	GATATCCAAATGGCTTATAGGCTCTTGAGGGTGAAGACAAATCCCTTGAGTTGGAAGG	660
Db	601	GATATCCAAATGGCTTATAGGCTCTTGAGGGTGAAGACAAATCCCTTGAGTTGGAAGG	660
Qy	661	CCAAAGCAAGGAATGTTGGTGTGTGGGTGGTGGATGTAACATGGAACCCGAGGATGCGAGG	720
Db	661	CCAAAGCAAGGAATGTTGGTGTGTGGGTGGTGGATGTAACATGGAACCCGAGGATGCGAGG	720
Qy	721	GTCGAAGTATTTACAAGACCTCGAAGAGATTTGAATTAATTCACACGATAGGCAAGCATTC	780
Db	721	GTCGAAGTATTTACAAGACCTCGAAGAGATTTGAATTAATTCACACGATAGGCAAGCATTC	780
Qy	781	GGAGATACCTGMAAGATTAATAATTCGATCCCAACATATCTATCTGTAAATTTATCTT	840
Db	781	GGAGATACCTGMAAGATTAATAATTCGATCCCAACATATCTATCTGTAAATTTATCTT	840
Qy	841	TCATTTGAAATCAATTCACAAAGATATCAACACAGAAAGGCTATCAATGATGATTTTG	900
Db	841	TCATTTGAAATCAATTCACAAAGATATCAACACAGAAAGGCTATCAATGATGATTTTG	900
Qy	901	GCGGTTCATATACCTGTGTCTGTGGGTGCATCTGAGGAAAACTATGAGAAATTTATCCCA	960
Db	901	GCGGTTCATATACCTGTGTCTGTGGGTGCATCTGAGGAAAACTATGAGAAATTTATCCCA	960
Qy	961	GCTGATTCATCTATCATGTGTGGAAGATTTTAACCTCCACGTAAGTATGACAAATATCTG	1020
Db	961	GCTGATTCATCTATCATGTGTGGAAGATTTTAACCTCCACGTAAGTATGACAAATATCTG	1020
Qy	1021	AAGGAAGTATGACAAAAACAAATAGTGTACCTTATGTTACTTTTACTGAGAAAGATTTT	1080
Db	1021	AAGGAAGTATGACAAAAACAAATAGTGTACCTTATGTTACTTTTACTGAGAAAGATTTT	1080
Qy	1081	ACTGTAACCTTACACAGGTTTGGGAATCATCATGCACTGCTGGCAATGGGACACATGATTA	1140
Db	1081	ACTGTAACCTTACACAGGTTTGGGAATCATCATGCACTGCTGGCAATGGGACACATGATTA	1140
Qy	1141	AGGCATCAAAATATATAGTCTGTGGTATTTAGAGAAATGGTTTGGATTAATAAGTGC	1200
Db	1141	AGGCATCAAAATATATAGTCTGTGGTATTTAGAGAAATGGTTTGGATTAATAAGTGC	1200
Qy	1201	CATCATGTGTACAGTGAATAAAGATTAATGATGACATTAATCAAGTTTTCAGATATAGAGA	1260
Db	1201	CATCATGTGTACAGTGAATAAAGATTAATGATGACATTAATCAAGTTTTCAGATATAGAGA	1260
Qy	1261	AACACGCTGATTTGGGGACCGTTTAAATTTTCCCTGCGCCCTCGGAGAGACATGATAT	1320
Db	1261	AACACGCTGATTTGGGGACCGTTTAAATTTTCCCTGCGCCCTCGGAGAGACATGATAT	1320
Qy	1321	TTTGGGTGAATTTTAAAGATCAAGATAGGAATCACTCAATATTTGGTTTAAATATCC	1380
Db	1321	TTTGGGTGAATTTTAAAGATCAAGATAGGAATCACTCAATATTTGGTTTAAATATCC	1380
Qy	1381	TGTATATATGTGATTAACGAGCACTGGAAATTAATTTATTTGTCACTCTATTTAATAACAT	1440
Db	1381	TGTATATATGTGATTAACGAGCACTGGAAATTAATTTATTTATTTGTCACTCTATTTAATAACAT	1440
Qy	1441	TGTTTTTACATTTTATAGTGACTGTAAAGTAAATTTATGATTTACTGTTTCAATCA	1500
Db	1441	TGTTTTTACATTTTATAGTGACTGTAAAGTAAATTTATGATTTACTGTTTCAATCA	1500

```

QY 1501 ATCGATCTTAAATCTATTTGGGAGATAAAAATACATATCCATAAATATGAGGAGCTTTT 1500
DB 1501 ATTCAGATCTTAAATCTATTTGGGAGATAAAAATACATATCCATAAATATGAGGAGCTTTT 1560
QY 1561 TGGCAATCTATTAATAGATAGATCTGATATATGCTATACAAAGAAATTTTGTG 1620
DB 1561 TGGCAATCTATTAATAGATAGATCTGATATATGCTATACAAAGAAATTTTGTG 1620
QY 1621 AATCAGTAAATCTTAAATCTTGAATCAAAAGTTGACACTGACACATGATAGAG 1680
DB 1621 AATCAGTAAATCTTAAATCTTGAATCAAAAGTTGACACTGACACATGATAGAG 1680
QY 1681 TTTTACTTGTACCATTTTGTGAAGGGAGCTAATTTGATGTAGCAATTAAGAAATAT 1740
DB 1681 TTTTACTTGTACCATTTTGTGAAGGGAGCTAATTTGATGTAGCAATTAAGAAATAT 1740
QY 1741 GAAGCAGAAATGTATATAGCTTCACATTTATTTTATAGTTTGTACCAATGTAGA 1800
DB 1741 GAAGCAGAAATGTATATAGCTTCACATTTATTTTATAGTTTGTACCAATGTAGA 1800
QY 1801 TTCCCAATATTTTAACTTAAGAAACAAAGATTTTCAGATTAACATTTATGTGTTT 1860
DB 1801 TTCCCAATATTTTAACTTAAGAAACAAAGATTTTCAGATTAACATTTATGTGTTT 1860
QY 1861 AGGAGATTTGATTTGTTTGTGGGTTTGTAGTCCCTAATATACCTTAAGAGAACTTT 1920
DB 1861 AGGAGATTTGATTTGTTTGTGGGTTTGTAGTCCCTAATATACCTTAAGAGAACTTT 1920
QY 1921 TGAATAGTCCCTTCAAAATCATGAGATACCTATTTTAAATTAATTTTCTCT 1980
DB 1921 TGAATAGTCCCTTCAAAATCATGAGATACCTATTTTAAATTAATTTTCTCT 1980
QY 1981 TATAACATTAATTAATATGATATCTCAATATCAAAATTAATGTATTAATTTATACATTTAT 2040
DB 1981 TATAACATTAATTAATATGATATCTCAATATCAAAATTAATGTATTAATTTATACATTTAT 2040
QY 2041 AGAAGCTAGTACTGATTCATTCATTTGAATATTTTTCACAAGCAGTCATATTTAACTTT 2100
DB 2041 AGAAGCTAGTACTGATTCATTCATTTGAATATTTTTCACAAGCAGTCATATTTAACTTT 2100
QY 2101 TATTTACTGATTTGTTATTAATTAAGATTAACCTGAAAAAATTTTAAAAA 2160
DB 2101 TATTTACTGATTTGTTATTAATTAAGATTAACCTGAAAAAATTTTAAAAA 2160
QY 2161 AAAAAAAAAA 2170
DB 2161 AAAAAAAAAA 2170

RESULT 2
AAZ92672
ID AAZ92672 standard; cDNA; 2036 BP.
XX
XX AAZ92672;
XX
XX 05-JUN-2000 (first entry)
XX
XX Murine alpha-1,3-fucosyltransferase gene exon 1 DNA, SEQ ID NO:30.
XX
XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
XX brain; kidney; recombinant expression; transgenic animal; knockout
XX animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
XX treatment; cancer; murine; mouse; ss.
OS
XX Mus sp.
XX
XX WO200006708-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-JP04092.
XX
XX 29-JUL-1998; 98JP-0213823.

```

```

XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Narimatsu H, Kudo T, Sasaki K;
XX
DR WPI; 2000-183120/16.
XX
PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer.
XX
PS Example 8; Page 162-164; 172pp; Japanese.
XX
CC The invention relates to a novel alpha-1,3-fucosyltransferase which
CC transfers a fucose moiety to galactosyl-beta-1,4-N-acetylgalucosamine
CC (generating the Lewis x or y epitope). It does not transfer a fucose
CC moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylgalucosamine
CC and therefore does not generate the sialyl-Lewis x epitope. The
CC invention also relates to DNA sequences encoding alpha-1,3-
CC fucosyltransferase and expression vectors and host cells comprising
CC these DNA sequences. The invention additionally encompasses the
CC preparation of alpha-1,3-fucosyltransferase via the culture of
CC transformed cells or by expression of the protein in a transgenic animal;
CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
CC screening potential inhibitors or potentiators of
CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
CC compounds having fucose-containing sugar chains by use of the protein;
CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
CC FUC-TIV and is expressed mainly in brain and kidney tissues.
CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
CC potentiators and inhibitors may be used for the treatment and diagnosis
CC of diseases of the brain and kidney, and of cancers. They may be used for
CC the identification of substances which affect the activity or expression
CC of alpha-1,3-fucosyltransferase; such substances may be used
CC therapeutically. The knockout animals can be used to study the mechanisms
CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
CC AAZ92645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
CC (AAZ90995), and sequences AAZ92670-292673 are murine
CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
CC amplification of cDNA ends).
XX
SQ Sequence 2036 BP; 666 A; 365 C; 348 G; 657 T; 0 other;

Query Match 93.8%; Score 2036; DB 21; Length 2036;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2036; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 GAAAAATTATGACATCAATCCAAAGCATTCCTGCCCCATTTCTAATGCTGTCATCA 166
DB 1 GAAAAATTATGACATCAATCCAAAGCATTCCTGCCCCATTTCTAATGCTGTCATCA 60
QY 167 TCCTGGGCTGCTCATGCGATGCTGCTCATTTATATCAAGCCACCAACAGCTGGGTCT 226
DB 61 TCCTGGGCTGCTCATGCGATGCTGCTCATTTATATCAAGCCACCAACAGCTGGGTCT 120
QY 227 TCAGTCCAAATGAGTCTGCAAGTTCTGCTGAAATGAAAAATTTCTTCTCCAAAAA 266
DB 121 TCAGTCCAAATGAGTCTGCAAGTTCTGCTGAAATGAAAAATTTCTTCTCCAAAAA 180
QY 287 CTGATTAATTTTAAAGAAATACATTCCTGTTGGGTATGGCCATTTGGGACAGACCTTTG 346
DB 181 CTGATTAATTTTAAAGAAATACATTCCTGTTGGGTATGGCCATTTGGGACAGACCTTTG 240
QY 347 ACCTTACATCTCTGCAAGCAATGTTCAATATTCGAAGGGTCCATCTCACAACAGACCGCT 406
DB 241 ACCTTACATCTCTGCAAGCAATGTTCAATATTCGAAGGGTCCATCTCACAACAGACCGCT 300
QY 407 CATTTGACAAATATCCATGCGGTCTGATTCACCATAGAGACATCAGCTGGATCTGA 466
DB 301 CATTTGACAAATATCCATGCGGTCTGATTCACCATAGAGACATCAGCTGGATCTGA 360
QY 467 CTAACTTACCTGACAGGCGCAGCCACCTTTCAGAAATGATTTGGATGAATTTAGAGT 526

```

```

Db 361 CTACCTTACCTGACGAGCCAGGCGACCCCTTTTCGAAATGATTTGATGATTTAGAGT 420
OY 527 CACCCACTCACACCCCCCAAAAAGATGGCATTTGAACACTTGTCAACTGACTTAACTT 586
Db 421 CACCCACTCACACCCCCCAAAAAGATGGCATTTGAACACTTGTCAACTGACTTAACTT 480
OY 587 ATGCGCGTGAATCAGATATCCAAAGTCCCTTATGCTTCTGACGCTGAGACAAATCCCT 646
Db 481 ATCCCGCTGATTCAGATATCCAAAGTCCCTTATGCTTCTGACGCTGAGACAAATCCCT 540
OY 647 TTGTGTTGAAGTCCCAAGGAGAGATGTTGCTGCTGGGTGGATTAAGTGAAC 706
Db 541 TTGTGTTGAAGTCCCAAGGAGAGATGTTGCTGCTGGGTGGATTAAGTGAAC 600
OY 707 CTGAGCATGGCCAGCTCAGCTTTTCAACGAGCTCAGCAAGAGTATGAAATCCACACT 766
Db 601 CTGAGCATGGCCAGGCTCAGGATATTCACAGGCTCAGCAAGAGTATGAAATCCACACT 680
OY 767 ATGCGCAAGCATTCGAGATATCGTGAACGATTAATAATCTGATCCACCATATCTACT 826
Db 661 ATGCGCAAGCATTCGAGATATCGTGAACGATTAATAATCTGATCCACCATATCTACT 720
OY 827 GTAAATTTTATCTTTCATTTGAAATCTCAATTCACAAAGATTAATCAGACAGAAAGCT 886
Db 721 GTAAATTTTATCTTTCATTTGAAATCTCAATTCACAAAGATTAATCAGACAGAAAGCT 780
OY 887 ACAATGCATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946
Db 781 ACAATGCATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
OY 947 AGAATTTATTTCCAGCTGATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 1006
Db 841 AGAATTTATTTCCAGCTGATTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
OY 1007 TAGCAAAATATCTGAAAGAGTTCACAAAACAAATAAGTTCACCTTACTTACTTACT 1066
Db 901 TAGCAAAATATCTGAAAGAGTTCACAAAACAAATAAGTTCACCTTACTTACTTACT 960
OY 1067 GGAGAAAGAGATTTTACTGTAACCTACACAGCTTTGGGATTCATTCATTCATTCATTC 1126
Db 961 GGAGAAAGAGATTTTACTGTAACCTACACAGCTTTGGGATTCATTCATTCATTCATTC 1020
OY 1127 GCGACCATGTAATAAGGCTCAGAAATTAAGTCTGTGGTAAATTAAGAAATGCTTTT 1186
Db 1021 GCGACCATGTAATAAGGCTCAGAAATTAAGTCTGTGGTAAATTAAGAAATGCTTTT 1080
OY 1187 GGAATTAAGTGTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1246
Db 1081 GGAATTAAGTGTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1140
OY 1247 TCAGGATTAAGAAACACGCTGATTTGGGCGACCTTTAATTTCCGCGCCCTG 1306
Db 1141 TCAGGATTAAGAAACACGCTGATTTGGGCGACCTTTAATTTCCGCGCCCTG 1200
OY 1307 AGGACCATGTAATTTTGGTGAATTTTAAAGATTCAGAAATTAAGTGAATTCATTCATTC 1366
Db 1201 AGGACCATGTAATTTTGGTGAATTTTAAAGATTCAGAAATTAAGTGAATTCATTCATTC 1260
OY 1367 GTTTTAAATTAATTCCTGTATATGTGATTAAGGACCTGGAATTAATTTATTTGTCACTC 1426
Db 1261 GTTTTAAATTAATTCCTGTATATGTGATTAAGGACCTGGAATTAATTTATTTGTCACTC 1320
OY 1427 TCATTTTAAACATTCCTTTTACATTTTATAGTGAATTCAGAAATTAATTTATTTGTCACTC 1486
Db 1321 TCATTTTAAACATTCCTTTTACATTTTATAGTGAATTCAGAAATTAATTTATTTGTCACTC 1380
OY 1487 TTGTTTCTACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1546
Db 1381 TTGTTTCTACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1440
OY 1547 TATGAGGAGCTTTTCCCAAGTATTAATAAGTATAGATTCAGCTGTAATAAGTCAACA 1606

```

```

Db 1441 TATGAGGAGCTTTTCCCAAGTATTAATAAGTATAGATTCAGCTGTAATAAGTCAACA 1500
OY 1607 AAGAAGATTAATTCATCACTTAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1666
Db 1501 AAGAAGATTAATTCATCACTTAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1560
OY 1667 GACAAGTATGAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1726
Db 1561 GACAAGTATGAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1620
OY 1727 AATAAGAAAGTATGAAGCAAGAAATTAATTAAGTATTCATTCATTCATTCATTCATTCAT 1786
Db 1621 AATAAGAAAGTATGAAGCAAGAAATTAATTAAGTATTCATTCATTCATTCATTCATTCAT 1680
OY 1787 TTAGCAAAATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1846
Db 1681 TTAGCAAAATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1740
OY 1847 AATTATATGTTTATGAGGAGATTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1906
Db 1741 AATTATATGTTTATGAGGAGATTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1800
OY 1907 CTAGAGAAATCTTTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1966
Db 1801 CTAGAGAAATCTTTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1860
OY 1967 AATATATTTCTCTTATTAACATTAATAAATTAATTAATTAATTAATTAATTAATTAATTA 2026
Db 1861 AATATATTTCTCTTATTAACATTAATAAATTAATTAATTAATTAATTAATTAATTAATTA 1920
OY 2027 AATTATACATTTATTAAGGCTATGATTCATTCATTCATTCATTCATTCATTCATTCATTC 2086
Db 1921 AATTATACATTTATTAAGGCTATGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1980
OY 2087 CCATATTTAAGTATTTTACTGATATGTTTAAATTAATTAATTAATTAATTAATTAATTA 2142
Db 1981 CCATATTTAAGTATTTTACTGATATGTTTAAATTAATTAATTAATTAATTAATTAATTA 2036

RESULT 3
AA292647
ID AA292647 standard; cDNA; 2676 BP.
XX
AC AA292647;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.
XX
KW Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW brain; kidney; recombinant expression; transgenic animal; knockout
KW animal; FUC-IV; drug screening; inhibitor; potentiator; diagnosis;
KW treatment; cancer; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 295..1374
FT /tag= a
FT /product= "Human alpha-1,3-fucosyltransferase"
XX
XX MO200006708-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99MO-JP04092.
XX
XX 29-JUL-1998; 98JP-0213823.
XX
XX (K10M ) KYOMA HAKKO KOGIO KK.
XX
XX Narimatsu H, Kudo T, Sasaki K;
XX

```


Qy	1550	GAGGAGCCTTTTCCCAAGATTTTAAAGATGATGATACCTGTTATATCTTACCAAG	1609
Db	1747	GA-AAAATTTTCCATAGATATTCAAATGCTGTGTTCCACCTTTGCCATCTTAACAGAG	1805
Qy	1610	AAGAATTATGTAATCATCTTAATCTTAACTCTTGACTTCAAGTTGACACTACTGAC	1668
Db	1806	AGAGA--CATGTGGGATGATCTTACCTCTTTGACCTCTTAAGATATAGAAGTGTAT	1863
Qy	1670	AACGT-----ATAGAGTTTTACTGTGACCAATTTGTGAGAGGAGCTAA	1715
Db	1864	AACCTCTCTATTTGATGATCTATTTTCTTTACCGTTTATACATCTTTGTGAAAGTGAATTA	1922
Qy	1716	TGATGTAGCAAAATAGAAATATGACACAGAAATGT-----TATTAGAC	1761
Db	1924	TTCATGGAGTGATTAAGAAATATGATGAGCAGAACTGTCTATTACGAGACTATTAGAC	1983
Qy	1762	TTCACTTTCATTTTATTATGATGTGTTTACCAATGTGATTTCCAAATTTTAAAGTTAA	1821
Db	1984	TTCTCATTTTATTTTCACTTACCTATTTGACACTACTTATTCATGCTTTAAATTTAA	2043
Qy	1822	GAAACAAAGATTTTCAATTAACATTTTATGCTTTTGGAGATTTGATGTTGTT	1881
Db	2044	TTATTCAGATTTTATTTTAAATATCCAAATTTGTTGATTT-----TTTCAGC	2087
Qy	1882	TTGTGTGTTTATGTCCTATATATACCTAAGAAATCTTTGATAGTCCCTTCAATCA	1941
Db	2088	ACCTGTGGAGATTAATCCCAATTAATCTTTGAAATAC--TAAGACAGTGTCTTCTGACTG	2146
Qy	1942	TGAAGATTAACCATTAATTTTAAAAATTAATTTTCCCTTATTAATCAATAAAAAATGATA	2001
Db	2147	ATGACACTCATGTCTAATAATTAACAATATATTTCTCTAATTAACAAGAAAAA-----	2200
Qy	2002	TCTTCAAAATCAAAATTAATATGATTAATTAATACACTATTAAGACGTACTATAC-TCAT	2066
Db	2201	-TGAATCCTATAAATATTTATTTATTAATGGTGCTATTTATGACAAATGTTTAAATATGTAT	2255
Qy	2061	CATTGAAATATTTTTCACAAAGCATCATTTTAA	2096
Db	2260	CAATTTAAGATTTTCTTGAAGCCCTTAATTTTAA	2295

XX	RESULT 4
XX	AA293646
XX	ID AA292646 standard; CDNA: 2822 BP.
XX	AA292646;
XX	05-JUN-2000 (first entry)
XX	Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:4.
XX	
XX	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
XX	brain; Kidney; recombinant expression; transgenic animal; knockout
XX	animal; FUC-TV; drug screening; inhibitor; potentiator; diagnosis;
XX	treatment; cancer; human; ss.
OS	Homo sapiens.
XX	
XX	
PH	Key
FT	Location/Qualifiers
FT	CDS 1289..2368
FT	/*tag= a
FT	/product= "Human alpha-1,3-fucosyltransferase"
XX	
PN	W0200006708-A1.
XX	
PD	10-FEB-2000.
XX	
PF	29-JUL-1999; 99WO-JP04092.
XX	
PR	29-JUL-1998; 98JP-0213823.
XX	
XX	(RYOW) KYOWA HAKKO KOGYO KK.
PA	
XX	

PI Nariimatsu H, Kudo T, Sasaki K,
XX
XX WPI; 2000-183120/16.
DR P-PSDB; AAY80996.
XX
XX Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer. -
XX
XX
XX Claim 4; Page 134-142; 172pp; Japanese.

CC	The invention relates to a novel alpha-1,3-fucosyltransferase which
CC	(transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
CC	generating) the Lewis x or y epitope). It does not transfer a fucose
CC	moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
CC	and therefore does not generate the sialyl-Lewis x epitope. The
CC	invention also relates to DNA sequences encoding alpha-1,3-
CC	fucosyltransferase and expression vectors; and host cells comprising
CC	these DNA sequences. The invention additionally encompasses the
CC	preparation of alpha-1,3-fucosyltransferase via the culture of
CC	transformed cells or by expression of the protein in a transgenic animal;
CC	antibodies which recognise alpha-1,3-fucosyltransferase; methods for
CC	screening potential inhibitors or potentiators of
CC	alpha-1,3-fucosyltransferase activity or expression; the preparation of
CC	compounds having fucose-containing sugar chains by use of the protein;
CC	and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
CC	Alpha-1,3-fucosyltransferase has a similar substrate range to the known
CC	FUC-IV and its expressed mainly in brain and kidney tissues.
CC	Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
CC	potentiators and inhibitors may be used for the treatment and diagnosis
CC	of diseases of the brain and kidney, and of cancers. They may be used for
CC	the identification of substances which affect the activity or expression
CC	of alpha-1,3-fucosyltransferase; such substances may be used
CC	therapeutically. The knockout animals can be used to study the mechanisms
CC	of action and expression of alpha-1,3-fucosyltransferase. Sequences
CC	AAG92646 and AAG92647 represent cDNAs encoding human
CC	alpha-1,3-fucosyltransferase (AA95096).
CC	XX
SQ	Sequence 2822 BP; 918 A; 485 C; 477 G; 942 T; 0 other;
Query Match	50.3% Score 1092.2; DB 21; Length 2822;
Best Local Similarity	85.8%; Pred. No. 2,4e-173;
Matches 1309; Conservative	0; Mismatches 198; Indels 18; Gaps 8;
OY	107 GAAAAATTATGACATCAACAATCCAAGCAGTCTTCGCCATTCTTAATCGTGCATCA 166
DB	1281 GAAAATAATTGACATCAACATCCAAAGGAATTTCTGGCCATTTTAAATGTGTGCATCA 1340
OY	167 TCCTGGGCGCTTCATGGCAGTGTCTCTCATTTAATCAAGCCCAACAGCTGGGTCT 226
DB	1341 TCCTGGGCGTTCATGGCAGTGTCTCTCATTTAATCAAGCCCAACAGCTGGGTCT 1400
OY	227 TCAGATCCAAATGGAGCTGTGCAAGTTCTGTGCTGTAAGAAATGAAATTTCTTCTCCACAAAA 286
DB	1401 TCAGATCCAAATGGAGATCCAGCAGCTGTGCTGTAAGAAATGAAATTTCTTCTCCACAAAA 1460
OY	287 CTGATTAATTTTTAACCAAAATACCATTCTGGTTGGTAGATGGCCATTGTGGCGAGACTTGG 346
DB	1461 CTGATTAATTTTAAATGAACACTAATTTCTGTGTGGGTGTGGCCATTTGGCGAGACTTTC 1520
OY	347 ACCTTCACATCCGCCAAGCAAGCATTTCAATATCCAAAGGTGCGATCTCACAAAGAGCGCT 406
DB	1521 ACCATTCATCTCGCCACAGCAATGTCTACATCTCCAAAGATGCGATCTCTACAAAGGACGCTT 1580
OY	407 CATTTGACAACAATCCCATCGGCTCGCTGATTCACCATAGACATCAGCTGGATCTGA 466
DB	1581 CACTGTACAACAAAATCCCATGCAAGTTCTGATCCATCACCGAGACATCACTTGGATCTGA 1640
OY	467 CTAACTTACCTAGCAGCGCCAGCGCACCCCTTTCAGAATATGGATTTGGAGATTAGAGT 526
DB	1641 CAATTTTACCTAGCAAGCTAGCGCACCCCTTTCAGAATATGGATTTGGAGATTGGAGT 1700
OY	527 CACCAACTACAGCCCCCAAAAGATGGGATTTGAGACATTTSTTCAACGTAATCACTT 586


```

Db 1701 CACCACTCACACTCCCAAGAGTGCATTGAGCAGCTGTTAACTGACCTGACCT 1760
Oy 587 ATGCGGATGATCATATATCAAGTGCCTTATGGCTTCTTGACGGTGAACCAATCCCT 646
Db 1761 ACCGCCGTATTCAGATATTCAGAGTCCCTTATGGCTTCTTGACGGTGAACCAATCCCT 1820
Oy 647 TTGTGTTTGAAGTCCCAAGCAAGAGAGTGTGTGTGCTGGTGTGTGATTAAGTGAAC 706
Db 1821 TCGTGTGTTGAGTGCACCAAGCAAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1880
Oy 707 CTGAGCATGCTCCAGGCTCAAGTATTCACAGAGCTCAGCAGAGTATTCAGATCCACCT 766
Db 1881 CTGAGCATGCTCCAGAGTCAAGTATTCAGATCCACCTCAGCAGAGTATTCAGATCCACCT 1940
Oy 767 ATGGCCAAAGCATTCGAGAGTACGTGAACGATTAATAATGATTCACCATATCTACT 826
Db 1941 ACGGCCAAGCATTTGAGAGATATTCATATGATTAATAATGATTCACCATATCTACT 2000
Oy 827 GTAAATTTTATCTTCATTTGAAAACCTCAATTCACAAAGATTACATCAGCAAGAAAGCTCT 886
Db 2001 GTAAATTTTATCTTCATTTGAAAACCTCAATTCACAAAGATTACATCAGCAAGAAAGCTCT 2060
Oy 887 ACAATGCAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
Db 2061 ACAATGCAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2120
Oy 947 ACAATTTATTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
Db 2121 ACAATTTATTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2180
Oy 1007 TAGCAAAATATCTGAGAGAGTTCACAAAACAAATAGTTTACCTTACTTAACT 1066
Db 2181 TAGCAAAATATCTGAGAGAGTTCACAAAACAAATAGTTTACCTTACTTAACT 2240
Oy 1067 GGAGAAAGGATTTTACTGTAACTACACGCTTTGGGAAATCAGATGATGCTGCTGCT 1126
Db 2241 GGAGAAAGGATTTTACTGTAACTACACGCTTTGGGAAATCAGATGATGCTGCTGCTGCT 2300
Oy 1127 GGGACCATTAAGAAAGCAAGCAATATTAAGTCTGTTGTTTAAAGAAAGGTTT 1186
Db 2301 GGGACCATTAAGAAAGCAAGCAATATTAAGTCTGTTGTTTAAAGAAAGGTTT 2360
Oy 1187 GGAATTT-AAAGTGTCCATCATCTGTACA---GTGAAAAGATTAATGAGACATTAATCCA 1242
Db 2361 GGAATTTAAATTTTTCATCAGTGTGACACTGTGATTAATTTTGAATGAGATTCATCCA 2420
Oy 1243 GTTTTCAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1297
Db 2421 GTATTGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2480
Oy 1298 CCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1356
Db 2481 CCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2540
Oy 1357 ACTCAATTTGTTTAAATTAATCTGATATATGATTAATGAGAGAGAGAGAGAGAGAGAGAG 1415
Db 2541 ACTCAATTTGTTTAAATTAATCTGATATATGATTAATGAGAGAGAGAGAGAGAGAGAGAG 2600
Oy 1416 ATTTGTCATCTCATTTTATTAATGATTTTAA---TTTAAAGTGCATTAAGAGAGAGAG 1472
Db 2601 ATTTGTCATCTCATTTTATTAATGATTTTAA---TTTAAAGTGCATTAAGAGAGAGAG 2660
Oy 1473 ATTTATGATTT---TACTGTTTCAATCATCATGATCTTTAACTTTTGGGAGATTA 1529
Db 2661 ACTCTGTGTTGATTTATTTGTTTCCACAGTGCATGATGCTGTTTAAATCTATTTGGGAGAG 2720
Oy 1530 AATATCATATCTCAAAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1589
Db 2721 AATATGAGATCTTAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTA 2779
Oy 1590 CTGTGATTAATGCTTAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1614

```

```

Db 2780 TTTCGATCTATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2804

RESULT 5
AA159505
ID AA159505 standard; cDNA; 1814 BP.
XX
XX AA159505;
AC
XX
XX 06-MAY-1997 (first entry)
DT
XX
XX Murine myeloid-lineage alpha-(1,3)-fucosyltransferase cDNA.
DE
XX
XX Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM;
KW septic shock; septicemia; therapy; ss.
XX
XX Mus sp.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 325..1353
FT /tag= a
PN
XX MO9640881-A1.
XX
XX 19-DEC-1996.
PD
XX
XX 08-MAY-1996; 96MO-US06427.
PF
XX 07-JUN-1995; 95US-0483151.
PR
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA
XX
XX Holgersson J, Seed B;
PI
XX
XX WPL 1997-108639/10.
DR
XX P-PSDB; AAM11820.
PR
XX
XX New murine alpha-(1,3)-fucosyltransferase - for fucosylating an
PT antibody to protect mammals against e.g. septic shock or septicemia
PS
XX
XX Claim 6: Fig 3; 58pp; English.
XX
XX
XX A cDNA clone (AA159505) codes for murine myeloid lineage alpha-
CC (1,3)-fucosyltransferase (AAM11820), an enzyme which has a relatively
CC strict substrate requirement for sialylated N-acetylglucosamine,
CC which can account for the presence of the sialyl-Lex epitope on
CC murine cells, and which is more effective than Fuc-TIV in support
CC of E-selectin-mediated COS cell adhesion. It was isolated from a
CC murine myeloid 32D c13 cell cDNA library by identification of a
CC clone capable of directing the expression of sialyl-Lex
CC determinants. Transformed host cells (e.g. 32D c13 or human 293
CC cells) expressing the alpha-(1,3)-fucosyltransferase, and pref-
CC erentially fucosyltransferase such as human Fuc-TIV (see also AA159506),
CC can be used to fucosylate an antibody (e.g. IgG or IgM) for use in
CC protecting an animal against an adverse immune reaction, esp. septic
CC shock or septicemia.
XX
XX
XX Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other:

Query Match 5.9%; Score 129; DB 18; Length 1814;
Best Local Similarity 52.3%; Pred. No. 6.9e-13;
Matches 335; Conservative 0; Mismatches 300; Indels 6; Gaps 2;
Oy 383 GGTGCATCTCAACACAGCCGCTATGTTACACAAATCCATCCGCTGATTCACC 442
Db 548 GGTGCATCTCAACACAGCCGCTATGTTACACAAATCCATCCGCTGATTCACC 607
Oy 443 ATAGAGACATCAGCTGGAGATCTGACTAATCTACCTCAGCAGGCGCAGCCCTTTCAGA 502
Db 608 ACCGAGACATCAGAAACCGGCAATCTCTACCCCTGAGCAGAGGCGCAGCAGCAGCAGCAGC 667
Oy 503 AATGATTTGATGAATTAATTAAGATCACCACACCCGCCCAAAAGAGTGATTAAGC 562

```

Tue Oct 8 10:18:16 2002

[illegible]

us-09-744-748-3.rng

FT FT 1960..2065
FT Intron /*tag=9
FT exon /number=9
FT 2067..3552
FT /*tag=f
FT /number=4
FT /codon_start=2126..2128
FT /note="alternative ATG start codon"
PN MO9732889-A1.
XX 12-SEP-1997.
XX 07-MAR-1997; 97WO-US03184.
XX 08-MAR-1996; 96US-0613098.
XX (UNMT) UNTV MICHIGAN.
XX Gersten KM, Lowe JB, Natsuka S;
XX WPI: 1997-470535/43.
XX P-PSDB; AAW27138 and AAW26671.
XX Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g.
XX screening for inhibitors used as antiinflammatory agents
XX Claim 1; Fig 2; 91pp; English.
CC This DNA sequence comprises the mouse Puc-VIII gene encoding
CC alpha-fucosyltransferase. It was isolated from a 3T3 cell genomic
CC library by screening with a probe from the catalytic domain of the
CC human Lewis alpha(1,3/4) alpha-fucosyltransferase. The catalytic domain of the
CC that mutation of the mouse Puc-VIII locus identified (Puc-VIII).
CC with different Golgi-localized domains of alpha(1,3) fucosyltransferases
CC (see AAW27138 and AAW26671). The proteins and nucleic acids can
CC be used to construct animal cell lines able to post-translationally
CC modify oligosaccharides on cell surfaces, intracellularly or
CC production of oligo, to isolate reagents, to generate antibodies to
CC glycosyltransferases useful as diagnostic reagents, to screen for
CC that act as antiinhibitors and inhibitors, to generate reagents, to screen for
CC the fucosyltransferase locus and for in vitro synthesis of sialyl
CC Lewis x tetrasaccharide.
SQ Sequence 3594 bp; 731 A; 969 C; 1067 G; 827 T; 0 other.
Query Match 5.9%; Score 128.6; DB 18; Length 3594;
Best local similarity 52.5%; Pred. No. 8,4e-13;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2.
Db 2277 GCTGCGCTCTACAGACAGACCGCTCATGTGTACAAACATCCATCGCGTCTGATTCAC 442
Oy 443 ATAGAGACATGACGTGGATGTGACTAACCTACACACAGCCAGCCAGCCACCTTTACAC 2336
Db 2337 ACCGTAGGTGAACACCGGCAATCTCTCTACCCCTGACACAGAGACCCACAGGACAGC 2396
Oy 503 AATGCAATTTGGATGAATTTGGATGATACCCACATCCACACCCGCCCAAAAGATGACATTCGAC 2396
Db 2397 CTGTGGCTGTGGCCCTCATATGCAATGCGCCCAATTAATACCATGATGCTCCATGCTTCCGG 2456
Oy 563 ACTGTGTACCTGACTTAATTAATGCGCGGTATTCAGATATTCAGATGCGCTTATGCGT 2456
Db 2457 GCATCTTCAACATCGGTGTGAGCTATGCGGATGATGATGATGATGATGATGATGATGATG 622
Oy 623 TCTTGACGCTGACACAAATCCCTTGTGTGATGATGATGATGATGATGATGATGATGATGATG 2516
Db 2517 GCTTGACGCTCTCTCTGCGCC--ACATCCCACTACCGGCGCAAAACAGATGATGCTG 2573

QY	683	GCCTGGTTTGTGATCTACGGAACCCGTGACATGCGCCAGGGTCAAGTATTTCACAAAGACTCA	742
Db	2574	CCTGGGTGATCAGCAATTTCCAGGAGCGGCACGACGCTCAAAAGCTGTACCGCACTGG	2633
QY	743	GCAGAGATTGAAATCCACACCTATGCGCAAGCATTCGGAGAAATACGTGAAGATMAA	802
Db	2634	CCCCATCATCTGGCAGAGTGGATGTGTTCGGTCCGCGCAGCGGAGCGGCCCATATGGGCTAAAT	2693
QY	803	ATCTGATTTCCACACATATCTACTTTGAAATTTTATCTTTCAATTTGAAACCTAATTCACA	862
Db	2694	GTCCTGTGCGCCACTTTGGCCCGGCTACCGCTCTACTGTGCGCTTTTGAAGACTACAGCATC	2753
QY	863	AGATTACATCTACAGAAAAGCTCTA--CAATGCATTTTGGCTGGTTCAGTACCTGTG	919
Db	2754	GGGATCATCTACCTAGAAAGTTCTGGCGCAATGCCCCGTGGCGTGTGATCCCGTGG	2813
QY	920	TCTGGGTCATCTAGGGGAAAACATATGAGAAATATATTCACAGTATTCATTCATTCATG	979
Db	2814	CGCTGGGACCTCTCTGGGGCCACTCTACAGAGGCTTTGTGCACACAGATGCTTTGATACAG	2873
QY	980	TGGAAATTTTACTCTCTCCCACTGAGTTAGC	1010
Db	2874	TGGAGACTTCACTGCTGCGCGTGAACCTGGC	2904
RESULT 7			
AA056912			
ID	AA056912 standard; DNA; 1086 BP.		
XX	AA056912;		
XX			
DT	26-JUL-1994 (first entry)		
XX			
DE	Lewis blood group fucosyltransferase DNA.		
XX			
KW	Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;		
KW	surface; oligosaccharide; ss.		
XX			
PN	W09402616-A.		
XX			
PD	03-FEB-1994.		
XX			
PF	20-JUL-1993; 93WO-US06703.		
XX			
PR	20-JUL-1992; 92US-0914281.		
XX			
PA	(UNMI) UNITV MICHIGAN.		
XX			
PI	Lowe JB;		
XX			
DR	WPI; 1994-048874/06.		
XX			
PT	DNA fragment encoding a glycosyltransferase - can be used for in		
PT	vitro reactions to modify cell surface oligosaccharide(s) e.g.		
XX	blood gp. determinants, to protect against transplant rejection		
XX			
PS	Disclosure: Fig 6; 249pp; English.		
XX			
CC	The sequence is that encoding a Lewis blood group fucosyl		
CC	transferase (Fuc-TfII). The enzyme produced by the DNA may be		
CC	non glycosylated. This prevents premature loss of enzyme activity.		
CC	It can also be used in in vitro reactions to modify cell surface		
CC	oligosaccharide mols. e.g. blood group determinants.		
CC	See also AA056905-11.		
XX			
SQ	Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;		
Query Match	5.9%;	Score 128.2;	DB 15; Length 1086;
Best Local Similarity	53.0%;	Pred. No. 9.2e-13;	
Matches 374;	Conservative 0;	Mismatches 313;	Indels 18; Gaps 4;
QY	385	TGCCATCTCACAACGACCGCTCATTTGACAAACAATCCCATGCGGTCTGTATTCACCAT	444

Db	271	TCGCATCATCGCGCGAGCGAAGGTGTACCCACGAGACAGGATCATGTCGACACC	330
Qy	445	AGAGCATCATGAGCTGGGA---TCTGACTAATCTTACCTCAGCAGGCCAGCCACTTCGAC	501
Db	331	TGGGATATCATGTCCACACCCCTMAGTCAAGCGCTCCACCCCTCCCGAGGCGCCAGGGGACG	390
Qy	502	AAATGATTTGGATGAAATTTAGAGTCAACCACTCACCTCACCCCCCAAAAGAGTGGATTTGAA	561
Db	391	CCCTGGATCTGGTTTAACTTTGGAGCCACCCCCCTTAAGTGCACAGCACTGGAAAGCCCTGGAC	450
Qy	562	CACCTGTTCAACCTGACTCTAATCTTATGCGCGGTGATTCAGATATCCAAAGTGCCTTATGGC	621
Db	451	AATATCTTCAATCTCAACCATGTCTACCGCAGCGACTCGCATGTCTTGACGCCCTTAAGGCC	510
Qy	622	TTCTTGAGGGTGAAGCACAATCCCTTTG-----TGTTTGAAGTGCACGACAAGAG	672
Db	511	TGGCTGGAGGCCCTGGTTCGGCGGAGCCTCCACACCCGCTCAACCTCTCGGCCAAGGCC	570
Qy	673	AAGTGTGATGCTGGGTGTGTGAGTAACTTGGAAACCCCTGAGCATGCAGGCGTCAAGTATTAC	732
Db	571	GAGCTGTGTGCTGGTGGCGGTGTCTCAACTGGAAGCCGGACTAGCCAGGGTCCGTACTAC	630
Qy	733	AACGAGCTCAGCAGAGATTTGAAATCCACACCTATGSGCAGCAATTCGGAAATACGTG	792
Db	631	CAGAGCCTGAGAGGCTCATCTCAAGGTGGACGTGTACGGAGCCTCCACA--ACCCCTG	687
Qy	793	AACGATTAATAATCTGATTTCCACCATATCTACTTAAATTTATCTTTCATTTTGAATAC	852
Db	688	CCCAAGGGGACCATATATGAGAGCGCTGTCCGGTTCAAAGTCTTACCTGGCCCTTGAGAAC	747
Qy	853	TCAATTCACMAAGATTATACATCAGAGAAAAGCTCT---ACAATGCAATTTTGTGGTGTCA	909
Db	748	TCCCTGCAACCCGCACTACATCATCCGAGAAAGCTGTGGAGAAAGCCCTGGAGGCTTGGGCC	807
Qy	910	GTACCTGTGTTCCTGGGTCTCATTTAGGAAAACTATGAAATTTATATCCAGCTGATTCA	969
Db	808	GTGCCCGTGGTGTGGGCCGACGAGAAAGCAACTACGAGAGTTCTCCACCCCGAGGCC	867
Qy	970	TTCAATTCATGTGGAGATTTTAACTCTCCAGTGAAGTATAGCAAAATATCTGACGAAGTT	1029
Db	868	TTCAATTCAGATGACGACTTCCAGAGCCCCAAGAACCTCGGCCGCTACTGCAAGAGCTG	927
Qy	1030	GACAAAAACATMACTGTACCTTAGTTACTTTAATCTGAGAAAG	1074
Db	928	GACAAGAGACCAAGCCCGCTACTGTGACTTGTGCTGGCGGAG	972
RESULT 8			
AAT80111			
XX	ID	AAT80111 standard; cDNA; 2042 BP.	
XX	AC	AAT80111;	
XX	DT	01-DEC-1997 (first entry)	
DE	DE	Fucosyltransferase III coding sequence.	
XX	XX	P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;	
KW	KW	calcium-dependent carbohydrate binding protein; platelet; endothelium;	
KW	KW	thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;	
KW	KW	inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;	
KW	KW	disseminated intravascular coagulation; rheumatoid arthritis; antibody;	
KW	KW	adult respiratory distress syndrome; tumour metastasis;	
KW	KW	fucosyltransferase III; acetylglucosaminyltransferase;	
XX	XX	autoimmune disorder; inflammatory disorder; ss.	
OS	OS	Homo sapiens.	
XX	XX		
PN	PN	W09706176-A2.	
XX	XX		
PD	PD	20-FEB-1997.	

Tue Oct 8 10:18:16 2002

us-09-744-748-3.ing

Page 10

```
PF 02-AUG-1996; 96WO-US12820.
XX
PR 17-MAY-1996; 96DS-0649802.
PR 03-APR-1996; 96DS-0510920.
PR 15-MAY-1996; 96DS-0017794.
XX
PA (OKLA ) UNIV OKLAHOMA STATE.
XX
PI Cummings RD, Meever RP, Moore KL;
XX
XX WPI: 1997-154206/14.
XX
XX New O-glycan(s) derived from P-selectin glycoprotein ligand
XX (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
XX for treating inflammatory or autoimmune disorders or tumours
XX
XX Disclosure: Page 80-81; 96pp; English.
XX
CC This sequence represents the coding sequence for the human
CC fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-
CC acetylglucosaminyltransferase encoded by AAT80112 are needed for the
CC expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in
CC mammalian cells. P-selectin is a calcium-dependent carbohydrate binding
CC protein expressed on the surface of activated platelets. PSGL-1 is
CC in response to thrombin and other activated platelets. PSGL-1 is a
CC P-selectin ligand and is calcium ion dependent by leukocytes. Binding of
CC treatment of the ligand with sialidase PSGL-1 is a homodimer. With a
CC highly extended intracellular domain, which is a feature of mucin-like
CC proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of
CC the invention (see AAT80112-W26188). PSGL-1 is also targeted by the
CC O-glycans, and O-glycosylated products of the invention. The products are
CC used for inhibiting binding of P-selectin to PSGL-1. The products are
CC resulting from ischaemia and reperfusion, bacterial infection, injury to
CC disseminated intravascular coagulation, adult respiratory distress
CC syndrome, tumour metastasis, rheumatoid arthritis, thrombocytopenia
CC in the production of antibodies and in diagnostic applications.
XX
XX Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;
XX
XX Query Match 5.9%; Score 128.2; DB 18; Length 2042;
XX Best local similarity 53.0%; Pred. No. 9,5e-13;
XX Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
XX
XX 385 TGCATCTCCACACAGACCGCTCATCTGATACACAAATCCATGCGCTGATTCACCAT 444
DB 343 TGCACATCACTCCGACCGCAAGGTGTACACAGCAGACAGCGTCACTGCGACAC 402
OY 445 AGACACATCACTGCGA--TGTGACTTAATCTTACCTGAGAGCCAGCCACCTTCAG 501
DB 403 TGGGATATCATGTCACACCTTAAGTCAAGCTCCACACTTCCCGAGGCGCAGGGGAG 462
OY 502 AATGATTTGATGATTAATTTAGATCAACCCAGCTACAGCCACCCCAAAAGATGCGATGAA 561
DB 463 CCGTGTGATCTGTTCACTTGAAGCCACCCCTTAATCTCCAGACCTGAGAGCCCTGGAC 522
OY 562 CACTTGTCAACCTGATCTTAATCTTGAAGCCGCTGATTCAGATTCAGCTTATGCG 621
DB 523 AGATCACTCAATCTCAACATGTCACCGCACTCCAGCAATCTTCACGCCCTCAGCG 582
OY 622 TTTTGAAGTGGCAAGCAAGAG 672
DB 583 TGGCTGAGCGCGGTGCGGCGCCAGCCACCCAGCCGCTCAACCTCTCGGCCAAGAC 642
OY 643 GAGCTGTGGTGGTGTGTGAGTAAGTGAACCCAGACCCAGATGAGGCTCAATATATTC 732
DB 733 AACGAGCTCAGCAGAGATTTAGTAATTCACACCTATGCGCAAGAGATTCGAGAAATCGTG 792
OY 703 CAGAGCCGCGAGGCTCATCTCAAGGTGAGCGTGTACGAGCGCTCCACA---AGCCCCG 759
XX
```

This sequence represents the coding sequence for the human fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-acetylglucosaminyltransferase encoded by AAT80112 are needed for the expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in mammalian cells. P-selectin is a calcium-dependent carbohydrate binding protein expressed on the surface of activated platelets. PSGL-1 is in response to thrombin and other activated platelets. PSGL-1 is a P-selectin ligand and is calcium ion dependent by leukocytes. Binding of treatment of the ligand with sialidase PSGL-1 is a homodimer. With a highly extended intracellular domain, which is a feature of mucin-like proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of the invention (see AAT80112-W26188). PSGL-1 is also targeted by the O-glycans, and O-glycosylated products of the invention. The products are used for inhibiting binding of P-selectin to PSGL-1. The products are resulting from ischaemia and reperfusion, bacterial infection, injury to disseminated intravascular coagulation, adult respiratory distress syndrome, tumour metastasis, rheumatoid arthritis, thrombocytopenia in the production of antibodies and in diagnostic applications.

Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

Query Match 5.9%; Score 128.2; DB 18; Length 2042; Best local similarity 53.0%; Pred. No. 9,5e-13; Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

```
OY 793 AACGATTAATAATCTGATCCACCAATCTACTGTAATTTATCTTCAATGAAAC 852
DB 760 CCCAGGGGACACATGATGAGAGAGCGTCCCGGACACATCTTACCTGAGAGAC 819
OY 853 TCAATTCACAAAGATTTAGATCAACAGAAAGCTCT---ACAATGCAATTTGGCTGTCA 909
DB 820 TCCCTGACACCCGACACATCATCACCAGAAAGCTGTGAGAGAACCCCTGAGGCTG 879
OY 910 GTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
DB 880 GTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
OY 970 TTTCAATTCATGTCGAGATTTTAACTCTCCAGTGAAGTTAGCAAAATATCTGAAGCAAGTT 1029
DB 940 TTTCAATTCATGTCGAGATTTTAACTCTCCAGTGAAGTTAGCAAAATATCTGAAGCAAGTT 1029
OY 1030 CACAAACAAATTAATGTTTAACTCTCCAGTGAAGTTAGCAAAATATCTGAAGCAAGTT 1029
DB 1000 GACAAAGACCCAGCCCGCTTACCTGATCTTAACTGAGGAAG 1074
XX
XX RESULT 9
XX ID AAO13330 standard; DNA; 2043 BP.
XX AC AAO13330;
XX
XX 07-NOV-1991 (first entry)
XX
XX Glycosyltransferase gene in clone pCDM7-alpha(1,3/1,4)FT.
XX
XX ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key
XX CDS
XX
XX Location/Qualifiers
XX FT 73..1158
XX FT /*lag= a
XX FT 2023..2034
XX FT /*lag= b
XX FT 1926..2003
XX FT repeat_region
XX FT 1783..1547
XX FT /*tag= d
XX FT 1885..1999
XX FT /*tag= e
XX FT note= "similar to human A1u sequences"
XX
XX MO9112340-A.
XX
XX 22-AUG-1991.
XX
XX 14-FEB-1991; 91WO-US00899.
XX
XX 12-DEC-1990; 90US-0627621.
XX 14-FEB-1990; 90US-0479858.
XX 14-FEB-1990; 90US-0480133.
XX
XX (UNMI ) UNIV OF MICHIGAN.
XX
XX Lowe JB;
XX
XX WPI: 1991-267151/36.
XX P-PSDB; AAR13749.
XX
XX Isolation of gene conveying post-translational characteristic -
XX e.g. the presence of soluble or membrane bound oligo or
XX polysaccharide or glycosyltransferase.
XX
XX Claim 11; Fig 1; 155pp; English.
XX
```

CC The DNA encodes a protein sequence capable of functioning as a
 CC GDP-Fuc:[beta-D-Gal(1,4/1,3)]-D-GlCNAc/(Glc)alpha(1,3/1,4)-
 CC fucosyltransferase. The sequence covers nucleotides 127 to 1083
 CC represents the functional protein. The DNA sequence can be used to
 CC construct animal cell lines with respect to posttranslational
 CC modification of the oligosaccharides on cell-surface, intracellular
 CC or secreted proteins or lipids by sub-terminal alpha(1,3) and alpha
 CC (1,4) fucose residues that represent the products of this enzyme.
 CC These can be used for the prodn. of diagnostics and therapeutics.
 CC See also AA013331-Q13333.

XX Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other;

Query Match

Best Local Similarity 53.0%; Score 128.2; DB 12; Length 2043;

Mismatches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 385 TGGCATCTCAGACAGACCGCTCATGTGACACAAATCCCATCGGTCCTGATTACCAT 444
 DB 343 TGGCATCTCAGACAGACCGCTCATGTGACACAAATCCCATCGGTCCTGATTACCAT 402
 QY 445 AGAGATCAATGAGTGGGA--TCTGACTAATCTACCTCAGAGCGCCACCCCTTTCAG 501
 DB 403 TGGGATATCATGTCCAAACCTTAAGTACAGCCTTCCACCTTCCCGAGGCGCGAGG 462
 QY 502 AATGATTTGATGATTTAGAGTCAACCCACTACACCCCGCCAAAAGATGGCATTTGA 561
 DB 463 CGCTGATCTGTTTCAACTTGGAGCCACCCCTTAAGTCCAGACCGTGAAGCCCTGGAG 522
 QY 562 CATTGTTCACCTGACTTAACTTATCGCCGATTCAGATATCCAGTCCCTTATGCG 621
 DB 523 AGATCTTCAATCTCAACATGCTCTACCGACGACTCCCATCTTCACGCTCCCTGCG 582
 QY 622 TTCTGACGGTGAACAATCCCTTGG-----TGTGTAAGTGGCAAGCAAGGAG 672
 DB 583 TGGCTGAGACCGGTGTCGGCCAGCCTGCCACCCACCTCAACTCTGCGCCAGAGAC 642
 QY 673 AAGTTGTGTGCTGGTGTGAGTAATGGAACCTGAGCAGTCCAGGGTCAATTTAC 732
 DB 643 GAGCTGTGGCTGGCGGCTGTCACACTGGAAAGCCGAGTCAACGAGGGTCCCTACTAC 702
 QY 733 AACGACTCAGCAAGATATGAAATCCACACTATGCGCAAGCATTCGGAATATCG 792
 DB 703 CAGAGCTCGAGGCTCATCTCAAGGTGAGGTGTACGAGGCTCCACA---AGCCCTG 759
 QY 793 AACGATAAAAATCTGATTCCTCCACCATATCTACTTGAATTTATCTTCAATTTGA 852
 DB 760 CCCAAGGGACCATGATGAGAGCGCTGTCCCGTACAAATTTCTACTGCGCTTCGAG 819
 QY 853 TCAATTCACAAAGATTTACATCACAGAAAGCTCT--ACAATGCAATTTTGGCTGTTCA 909
 DB 820 TCTTGCACCCCGACTACATACACGAGAAAGCTGTGAGAGAACGCTCGAGGCTGGGCC 879
 QY 910 GTACCTGTGTCTCTGGGTCATCTAGGAAAAGTATGAGATTTATTTCCAGTGTATCA 969
 DB 880 GTGCTCGTGTGTGTGGCCCGACAGAGCACTACAGAGGTGTCTTGCACCCGAGCGCC 939
 QY 970 TTCAATTCATGTGGAAGATTTTAATCTCCAGTGAAGTATGAGCAAAATATCTGAG 1029
 DB 940 TTCAATTCATGTGGAAGATTTTAATCTCCAGTGAAGTATGAGCAAAATATCTGAG 999
 QY 1030 GACAAAACAAATTAAGTTGACTTACTTACTTAACTTGAAGAAAG 1074
 DB 1000 GACAAAGACAGCGCCGCTACCTGAGACTTCTTGGTGGGGAG 1044

RESULT 10

AA056906

ID AA056906 standard; DNA; 2043 BP.

XX AA056906;

DT 26-JUL-1994 (first entry)

XX DNA encoding a glycosyltransferase.

DE Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
 XX surface; oligosaccharide; ss.

KW Homo sapiens.

OS Key Location/Qualifiers

FT CDS 72..1158

FT /tag- a

PN W09402616-A.

PD 03-FEB-1994.

XX 20-JUL-1993; 93WO-US06703.

XX 20-JUL-1992; 92OS-0914281.

PA (UNMI) UNITV MICHIGAN.

PI Lowe JB;

XX WPI: 1994-048874/06.

DR P-PSDB; AAR45934.

XX DNA fragment encoding a glycosyltransferase - can be used for in
 PT vitro reactions to modify cell surface oligosaccharide(s) e.g.
 PT blood gp. determinants, to protect against transplant rejection

PS Disclosure; Fig 1; 249pp; English.

XX The sequence is that encoding human glycosyl transferase. The enzyme
 CC produced by the DNA may be non glycosylated. This prevents premature
 CC loss of enzyme activity. It can also be used in in vitro reactions to
 CC modify cell surface oligosaccharide mols. e.g. blood group determinants.
 CC See also AA056905-12.

SO Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other;

Query Match

Best Local Similarity 53.0%; Score 128.2; DB 15; Length 2043;

Mismatches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 385 TGGCATCTCAGACAGACCGCTCATGTGACACAAATCCCATCGGTCCTGATTACCAT 444
 DB 343 TGGCATCTCAGACAGACCGCTCATGTGACACAAATCCCATCGGTCCTGATTACCAT 402
 QY 445 AGAGATCAATGAGTGGGA--TCTGACTAATCTACCTCAGAGCGCCACCCCTTTCAG 501
 DB 403 TGGGATATCATGTCCAAACCTTAAGTACAGCCTTCCACCTTCCCGAGGCGCGAGG 462
 QY 502 AATGATTTGATGATTTAGAGTCAACCCACTACACCCCGCCAAAAGATGGCATTTGA 561
 DB 463 CGCTGATCTGTTTCAACTTGGAGCCACCCCTTAAGTCCAGACCGTGAAGCCCTGGAG 522
 QY 562 CATTGTTCACCTGACTTAACTTATCGCCGATTCAGATATCCAGTCCCTTATGCG 621
 DB 523 AGATCTTCAATCTCAACATGCTCTACCGAGGACTCCGACATCTTCACGCTCCCTGCG 582
 QY 622 TTCTGACGGTGAACAATCCCTTGG-----TGTGTAAGTGGCAAGCAAGGAG 672
 DB 583 TGGCTGAGACCGGTGTCGGCCAGCCTGCCACCCACTCAACTCTGCGCCAAAGAC 642
 QY 673 AAGTTGTGTGCTGGTGTGAGTAATGGAACCTGAGCAGTCCAGGGTCAATTTAC 732
 DB 643 GAGCTGTGGCTGTGGCGGCTGTCACACTGGAAAGCCGAGTCAACCGAGGCTGCTACTAC 702
 QY 733 AACGACTCAGCAAGATATGAAATCCACACTATGCGCAAGCATTCGGAATATCG 792
 DB 703 CAGAGCTCGAGGCTCATCTCAAGGTGAGGTGTACGAGGCTCCACA---AGCCCTG 759

OY 793 AACGATTAATCTGATCCACCATCTACTTGTAAATTTTATCTTCAATTTGAAAC 852
 Db 760 CCCAAGGGACCATGATGAGACGCTCCCGGTACAACTTCTACTGCTGCGTGAAC 819
 OY 853 TCAATTCACAAAGATTACATCAGAAAAGCTCT--ACATGCAATTTTGGCTGTCA 909
 Db 820 TCTTTCACACCCGACTACATACGAGAACCTGTGAGAAAGCCCTGAGGCTGGGCC 879
 OY 910 GTACCTGTTCCTGGGTCATCAGGAAAATATGAAATATATCTCAGCTGATTC 969
 Db 880 GTGCCGTGTGTGGGCGCCACGAGAAAGAACTACAGAGCTTCTCTCCACCCGAGCC 939
 OY 970 TTCAATTCATGAGAAATTTAACTCTCCAGTGAATGCAAAATATCTCAGAAAGTT 1029
 Db 940 TTCAATTCATGAGAAATTTAACTCTCCAGTGAATGCAAAATATCTCAGAAAGTT 999
 OY 1030 GACAAAACAAATAGTTTACCTTACTTAACTGAGAAAG 1074
 Db 1000 GACAAAGACACAGCCCGCTACTGAGCTTCTGCTGCGGAG 1044

 RESULT 11
 ID AAT76769 standard: cDNA: 2043 BP.
 AC AAT76769;
 XX
 DT 15-SEP-1997 (first entry)
 DE Human alpha 1,3/4 fucosyltransferase cDNA.
 XX
 KW GDP-L-fucose:beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase;
 KW human; alpha 1,2 FT; transgene; transgenic mouse; animal model;
 KW intestinal adhesion; Helicobacter pylori infection; stomach;
 KW small intestine; gut; epithelial cell; surface receptor;
 KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
 KW gastric adenocarcinoma; Lewis antigen; fucosylation; ds.
 OS
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 104..1201
 FT /*tag= a
 XX
 PN US5652124-A.
 PD 29-APR-1997.
 XX 11-JUL-1994; 94US-0273411.
 PR 11-JUL-1994; 94US-0273411.
 XX (UNIW) UNIV WASHINGTON.
 PA Falk P, Gordon JI;
 PI
 XX MPI: 1997-258275/23.
 DR P-PSDB: AAW23806.
 XX
 PT Animal model for Helicobacter pylori infection - comprising
 PT transgenic mouse expressing human enzyme promoting intestinal
 PT adhesion
 XX
 PS Example 1: Columns 23-26; 24pp; English.
 XX
 CC A claimed transgenic mouse expresses, in its intestinal epithelial
 CC cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
 CC fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
 CC beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
 CC called alpha 1,3/4 FT). The enzyme is expressed under the
 CC control of a gut epithelial cell-specific promoter and Helicobacter
 CC pylori adheres to the transgenic cells. The transgenic mouse and
 CC intestinal epithelial cells from it are useful as models for screening

CC compounds for the ability to inhibit adhesion of H. pylori to gut
 CC epithelial cells. The present sequence encodes human alpha 1,3/4 FT
 CC and was published by Kukowska-Latallo et al., Genes Dev., 4:
 CC 1288-1303 (1990).
 XX
 SO Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other;
 Query Match 5.9%; Score 128.2; DB 18; Length 2043;
 Best Local Similarity 53.0%; Pred. No. 9.5e-13;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

 OY 385 TGGCATCTACAAACGACCGCTCATTTACACAAATCCATGCGTCTTACCAT 444
 Db 343 TGGCATCTACAAACGACCGCTCATTTACACAAATCCATGCGTCTTACCAT 402
 OY 445 AGAGACATCAGCTGGA--TCTGACTTACTTACCTCAGACGAGCCACCCCTTCA 501
 Db 403 TGGGATATCATGTCACACCCCTTACGCTCCACCTTCCCGAGGCGCAGGGGAG 462
 OY 502 AATGCAATTTGATGATTAATTTAGATCACCCTCAGACCCCAAAAGAGTGATGAA 561
 Db 463 CGCTGATCTGTTCACTTGAAGCCACCCCTTACCTGCAACACCTGGAAGCCCTGAG 522
 OY 562 CACTTGTCAACCTGACTCTTAACTTATGCGCGTATGAGATTCAGATTCAGTATGCG 621
 Db 523 AGATACTTCAATCTCAGCATCTCTTACCGAGGAGCTCCGACATCTTACCGCTTACGG 582
 OY 622 TTCTTGACGCTGAGCAAAATCCCTTTG-----TGTGAAAGTGCAGAGAGAG 672
 Db 583 TGGCTGAGACCGCTGCTGCGGACGCTGCGCACCACCCCTTACCTTCCGAGCAAGC 642
 OY 673 AAGTGTGCTGCTGCTGCTGATGATGATGAGACCTGAGCAATCCAGGGGTAAATAT 732
 Db 643 GACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
 OY 733 AAGCACTCAGCAAGATTTGAATTCACACCTATGCGCAAGCAATGAGAGATGAGT 792
 Db 703 CAGAGCTCAGAGCTCATCTCAAGTGTGAGCTGACGAGCTGCGCA--ACCCCTG 759
 OY 793 AACGATTAATCTGATCCACCATCTACTTGTAAATTTTATCTTCAATTTGAAAC 852
 Db 760 CCCAAGGGACCATGATGAGACGCTCCCGGTACAACTTCTACTGCTGCGTGAAC 819
 OY 853 TCAATTCACAAAGATTACATCAGAAAAGCTCT--ACATGCAATTTTGGCTGTCA 909
 Db 820 TCTTTCACACCCGACTACATACGAGAACCTGTGAGAAAGCCCTGAGGCTGGGCC 879
 OY 910 GTACCTGTTCCTGGGTCATCAGGAAAATATGAAATATATCTCAGCTGATTC 969
 Db 880 GTGCCGTGTGTGGGCGCCACGAGAAAGAACTACAGAGCTTCTCTCCACCCGAGCC 939
 OY 970 TTCAATTCATGAGAAATTTAACTCTCCAGTGAATGCAAAATATCTCAGAAAGTT 1029
 Db 940 TTCAATTCATGAGAAATTTAACTCTCCAGTGAATGCAAAATATCTCAGAAAGTT 999
 OY 1030 GACAAAACAAATAGTTTACCTTACTTAACTGAGAAAG 1074
 Db 1000 GACAAAGACACAGCCCGCTACTGAGCTTCTGCTGCGGAG 1044

 RESULT 12
 ID AAT61675 standard: cDNA: 2043 BP.
 AC AAT61675;
 XX
 DT 19-JUN-1997 (first entry)
 DE Human alpha (1,3/1,4)-fucosyltransferase (Lewis enzyme) cDNA.
 XX
 KW Alpha(1,3/1,4)-fucosyltransferase; Lewis enzyme; Fuc-TIII;
 KW glycosylation; oligosaccharide; blood group; ss.

```

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 73..1158
FT /*tag= a
FT
XX
XX MO9709421-A1.
XX
XX 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-US13816.
XX
XX 08-SEP-1995; 95US-0525058.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Legault DJ, Lowe JB:
XX
XX WPI: 1997-192897/17.
XX P-PSDB; AAM13638.
XX
XX New recombinant fucosyltransferase proteins - useful for modifying
XX cell surface oligosaccharide structures
XX
XX Example 3; Page 268-269; 329pp; English.
XX
XX A cDNA clone (AAT61675) codes for an enzyme (AAM13638) that functions as
XX a GDP-Fuc:(beta-D-Gal(1,4/1,3))-D-GlcNAc(Glc) alpha(1,3/1,4)-
XX fucosyltransferase (Fuc-TIII) which, when expressed by the cloned
XX cDNA sequence, functions within mammalian cells to generate de novo
XX expression of specific SSEA-1 (Lewis x), sialyl-Lewis x, Lewis a and
XX sialyl-Lewis structures. The clone was identified by transfecting
XX COS-1 cells with genomic DNA from human A431 cells, and isolating a
XX cloned cDNA that deid. expression of surface-localised SSEA-1
XX structures. The cDNA can be used to construct animal cell lines
XX with specific capabilities with respect to post-translational
XX modification of the oligosaccharides of expressed proteins or
XX lipids, to produce recombinant Fuc-TIII for use in oligosaccharide
XX prodn., to genotype individuals for the Lewis locus, and in the
XX prodn. of chimeric fucosyltransferases (see also AAM14512-32).
XX
XX
XX Sequence 2043 BP; 413 A; 646 C; 576 G; 408 T; 0 other;
XX
XX Query Match 5.9%; Score 128.2; DB 18; Length 2043;
XX Best Local Similarity 53.0%; Pred. No. 9.5e-13;
XX Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
XX
XX 385 TGGCATCTGCACACAGACCGCTCATTTGTACAAACAATCCCATGGCTCTGATTCACAT 444
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 343 TGGCATCTGCACACAGACCGCTCATTTGTACAAACAATCCCATGGCTCTGATTCACAT 402
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 445 AGAGACATCAGCTGGGA---TGTGACTACTTACTTACCTCAGCAGCGCCAGCCCTTTGAC 501
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 403 TGGGATATCATGTCCACACCTTAAGTACGCTTCCACCTTCCCGGCGGCGGAG 462
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 502 AATGATTTGGATGATTAATTAGATGACCCACTCACACCCCAAAAAGAGTGGCATTTGA 561
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 463 CGGTGATCTGGTTAACTTGGAGCCACCCCTTAAGTCTGACAGACTGGAAGCCCTGGAG 522
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 562 CACTTGTTAACCTGACTTAACCTTAATGCGCGGATTCAGATTCAGATTCCTTAATGCG 621
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 523 AGATACCTTCAATCTCACCAATGTCCTACCGCAGACTCCGACATCTTCACAGCCCTACGCG 582
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 622 TTTCTGACGCTGAGCACAATCCCTTTG-----TGTTGAAGTGGCAGCAGGAG 672
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 583 TGGCTGGAGACCGTGGTGGCGGCAAGCTGCCACACCAACCCCTCAACTCTTCGGCCAGACCC 642
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 673 AAGTGTGTGCTGGTGGTGGTGGTGAATGAAACCTGAGACATCCAGCGCTCAAGTATTAC 732
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 643 GAGCTGTGGCTGGGCGGTGTCACACTGGAAGCGGACTCAGCGGCTGGCTACTAC 702
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 733 TGGCATCTGCACACAGACCGCTCATTTGTACAAACAATCCCATGGCTCTGATTCACAT 792
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 703 CAGAGCCCTGCAGGCTCATCTCAAGGTGACGTGTACGAGCGCTCCACACA---AGCCCTGG 759
XX
XX 793 AACGATAAAATCTGATTCCTCCACCATATCTGTAATTTATCTTCATTTGAAGAAC 852
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 760 GCCAAGGAGACCATGATGAGACGCTGTCCCGGATACAAAGTTCTACTGGCTTCGAGAAC 819
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 853 TCAATTCACAAAGATTTACATACAGAAAAGCTCT---ACAATGCATTTTGGCTGGTTGA 909
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 820 TCCTTGACCCCGACTACATACACGGAAGCTGTGAGAGAACCCCTGGAGGCTGGGCC 879
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 910 GTACCTGTGTCTGGGTGCATCTAGGAAAACATAGAGATTAATTTCCAGCTGATTTCA 969
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 880 GTGCCCGGTGTCTGGGCCCGACAGAGCAACTACGAGAGTTCCTGCCACCCGACGCC 939
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 970 TTGATTCATGTGGAGATTTTAACTCTCCAGTAGTGTGCAAAATATCTGAAGAGTT 1029
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 940 TTCTTCACAGTGTGACGACTTCCAGAGCCCGGAGACTGGCCGGTACCTGAGGAGCTG 999
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1030 GACAAAAACAAATTAAGTTGACTTACTTACTTACTTACTTACTTACTTACTTACTT 1074
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1000 GACAAAGACCAACGCGCGTACCTGAGCTACTTTCGCTGGCGGAG 1044
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX RESULT 13
XX ID AAO56911 standard; DNA; 1316 BP.
XX
XX AC AAO56911;
XX
XX 26-JUL-1994 (first entry)
XX
XX DE DNA encoding a glycosyltransferase.
XX
XX KW glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
XX surface; oligosaccharide; se.
XX
XX OS Homo sapiens.
XX
XX PN WO9402616-A.
XX
XX PD 03-FEB-1994.
XX
XX PF 20-JUL-1993; 93WO-US06703.
XX
XX PR 20-JUL-1992; 92US-0914281.
XX
XX PA (UNMI ) UNIV MICHIGAN.
XX
XX PS Lowe JB:
XX
XX WPI: 1994-048874/06.
XX P-PSDB; AAR45939.
XX
XX DNA fragment encoding a glycosyltransferase - can be used for in
XX vitro reactions to modify cell surface oligosaccharide(s) e.g.
XX blood gp. determinants, to protect against transplant rejection
XX
XX PS Disclosure; Fig 6; 249pp; English.
XX
XX CC The sequence is that encoding human glycosyl transferase. The enzyme
XX CC produced by the DNA may be non glycosylated. This prevents premature
XX CC loss of enzyme activity. It can also be used in in vitro reactions to
XX CC modify cell surface oligosaccharide mols. e.g. blood group determinants.
XX CC See also AAO56905-12.
XX
XX
XX Sequence 1316 BP; 236 A; 456 C; 372 G; 252 T; 0 other;
XX
XX Query Match 5.8%; Score 126.6; DB 15; Length 1316;
XX Best Local Similarity 52.9%; Pred. No. 1.7e-12;
XX Matches 373; Conservative 0; Mismatches 314; Indels 18; Gaps 4;
XX
XX 385 TGGCATCTGCACACAGACCGCTCATTTGTACAAACAATCCCATGGCTCTGATTCACAT 444
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Tue Oct 8 10:18:16 2002

us-09-744-748-3.ing

Page 14

```

Db 430 TGCAGATGACGCTCCGACGCTGAGTGTACCCACAGGACCGGCTACGCTGACACAC
OY 445 AGAGATGACGCTGGAGTCTGA---CTAAGTACGCTGAGGACCGGACCGCTTTGAG 489
Db 490 TGGGATATCATGTACACCCGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 501
OY 502 AATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
Db 550 CCGCTGATGCTGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
OY 562 CACTGTTCACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 609
Db 610 GGAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 621
OY 622 TTTCTGAGGCTGAGCACAATTCCTTTG-----TGTGAAAGTGCACAGGAGAG 669
Db 670 TGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 729
OY 673 AAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
Db 730 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
OY 733 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
Db 790 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
OY 793 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
Db 847 CCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 852
OY 853 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
Db 907 TCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 909
OY 910 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Db 967 GTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
OY 970 TTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
Db 1027 TTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
OY 1030 GACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1074
Db 1087 GACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131

```

```

XX (DNMI) UNIV MICHIGAN.
PA Legault DJ, Lowe JB:
XX WPI: 1997-192897/17.
XX P-PSDB: AAM13642.
XX
PT New recombinant fucosyltransferase proteins - useful for modifying
PS cell surface oligosaccharide structures
XX
XX Example 5: Page 287-288; 329pp; English.
CC A DNA clone (AAT61679) codes for an enzyme (AAM13642) that functions as
CC a GDP-Fuc:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-fucosyltransferase
CC (Fuc-TV) which can utilize neutral type 1 oligosaccharide
CC precursors but which cannot efficiently utilize type I
CC glycoconjugates. It was identified in a human lambda phage
CC genomic DNA library using a probe obtd. from Lewis x enzyme cDNA (see
CC also AAT61675). The Fuc-TV enzyme, when expressed by the cloned DNA
CC sequence, functions within mammalian cells to generate de novo
CC expression of specific cell surface glycoconjugate structures that
CC are recognised by antibodies against SSEA-1, Lewis x or sialyl
CC Lewis x.
SQ
Sequence 1316 BP; 236 A; 456 C; 372 G; 253 T; 0 other:
Query Match 5.88; Score 126.6; DB 18; Length 1316;
Best Local Similarity 52.98; Pred. No. 1.7e-12;
Matches 373; Conservative 0; Mismatches 314; Indels 18; Gaps 4.
OY 385 TCCCATCTCACACAGACCCCTAATGTAACAAATCCATGCGGTCGATGACACAT 444
Db 430 TGCAGATGACGCTCCGACGCTGAGTGTACCCACAGGACCGGCTACGCTGACACAC 489
OY 445 AGAGATGACGCTGGAGTCTGA---CTAAGTACGCTGAGGACCGGACCGCTTTGAG 501
Db 490 TGGGATATCATGTACACCCGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGAC 549
OY 502 AATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
Db 550 CCGCTGATGCTGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
OY 562 CACTGTTCACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 609
Db 610 GGAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 621
OY 622 TTTCTGAGGCTGAGCACAATTCCTTTG-----TGTGAAAGTGCACAGGAGAG 669
Db 670 TGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 729
OY 673 AAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
Db 730 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
OY 733 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
Db 790 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
OY 793 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
Db 847 CCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
OY 853 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
Db 907 TCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
OY 910 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Db 967 GTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
OY 970 TTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 : Search time 1816.88 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-3

Perfect score: 2170
Sequence: 1 cttccctcgcctgggacgc.....aaaaaaaaaaaaaaaaaaaaa 2170

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estbta:*
2: em_estbhm:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	679.8	31.3	697	9	BB397507 BB397507
2	639	29.4	649	9	AV338075 AV338075
3	635.8	29.3	650	9	AV375511 AV375511
4	580.8	28.8	607	10	BB102760 BB102760
5	449.8	20.7	648	9	BB595164 BB595164
6	429	19.8	625	9	BB649967 BB649967
7	418.4	19.3	716	9	AV725071 AV725071
8	416.8	19.2	767	9	BB640911 BB640911
9	371.2	17.1	579	10	BM021081 BM021081
10	369.8	17.0	628	9	BB651409 BB651409
11	304.8	14.0	1101	12	CNS0595T CNS0595T
12	292.8	13.5	316	9	BB130253 BB130253
13	290.6	13.4	455	9	AI849151 UI-M-AJ1-
14	271.4	12.5	670	9	BB626126 BB626126
15	268	12.4	284	9	BB270301 BB270301
16	265.4	12.3	276	9	BB383634 BB383634
17	265	12.2	568	10	BM021317 BM021317

18	259.8	12.0	317	9	BB272545 BB272545
19	239.2	11.0	501	12	BA1073 BA1073
20	236.2	10.9	315	9	BB250973 BB250973
21	179.6	8.3	1100	12	CNS0579Z CNS0579Z
22	175.8	8.1	844	12	CNS01V8V CNS01V8V
23	172.2	7.9	199	9	BB398103 BB398103
24	171.4	7.9	570	10	BM082484 BM082484
25	169.8	7.8	661	9	BB651647 BB651647
26	169	7.6	367	9	BB843869 BB843869
27	164.8	7.3	887	12	CNS0200S CNS0200S
28	159.4	7.2	548	12	AQ72584 AQ72584
29	156.8	7.2	1016	12	CNS0429P CNS0429P
30	148.4	6.8	666	10	BB050300 BB050300
31	145.8	6.7	651	10	BB036855 BB036855
32	142	6.5	640	10	BB073227 BB073227
33	140.4	6.5	711	12	CNS034XF CNS034XF
34	127	5.9	533	10	BB604066 BB604066
35	126	5.8	674	10	BB605503 BB605503
36	125.4	5.8	322	9	BB498510 BB498510
37	115.6	5.3	751	9	BB260180 BB260180
38	115.6	5.3	971	12	CNS047IV CNS047IV
39	111.6	5.1	1101	12	CNS00EVL CNS00EVL
40	110.8	5.1	1023	12	CNS02ANG CNS02ANG
41	110.2	5.1	615	10	BB132681 BB132681
42	106.2	4.9	487	10	BB663396 BB663396
43	106.2	4.9	720	10	BB180380 BB180380
44	105.8	4.9	702	10	BB043578 BB043578
45	97.6	4.5	1101	12	CNS0021U CNS0021U

ALIGNMENTS

RESULT 1
LOCUS BB397507
DEFINITION BB397507 RIKEN full-length enriched, ES cells Mus musculus CDNA clone C330004C19 3' similar to AB015426 Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, mRNA sequence.
ACCESSION BB397507
VERSION BB397507.2 GI:16411808
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 697)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyata,T., Muramatsu,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 15, 2000 this sequence version replaced gi:9216903.
Contact: Yoshitake Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Ikawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

OY	121	TCACATCCAAAGCATCTTCGGCCATTTCTATTCGTGCATCATCTCGTGGCTTC	180
Db	316	TCACATCCAAAGCATCTTCGGCCATTTCTATTCGTGCATCATCTCGTGGCTTC	375
OY	181	ATGGCATCTGCGTCATTTATATCAAGCCACACAAAGCGTGGCTTCATAGTCAATGGAG	240
Db	376	ATGGCATCTGCGTCATTTATATCAAGCCACACAAAGCGTGGCTTCATAGTCAATGGAG	435
OY	241	TCGCGAATTCGTGTCGTGAATAAGAAAAATTTCTTCGCCAAAAACGTATATTTTAAC	300
Db	436	TCGCGAATTCGTGTCGTGAATAAGAAAAATTTCTTCGCCAAAAACGTATATTTTAAC	495
OY	301	GAATACCATCTTCGTGTTGGGTATAGGCAATTTGGGAGACCTTTGACCTTACATCCTGC	360
Db	496	GAATACCATCTTCGTGTTGGGTATAGGCAATTTGGGAGACCTTTGACCTTACATCCTGC	555
OY	361	CAGCAATGTTCAATATCCAAAGGTGCCATCTCAACAAGACCGCTCATGTAGACAAAA	420
Db	556	CAGCAATGTTCAATATCCAAAGGTGCCATCTCAACAAGACCGCTCATGTAGACAAAA	615
OY	421	TCGCATGGCGTCTGATTCAACATAGAGACATC	453
Db	616	TCGCATGGCGTCTGATTCAACATAGAGACATC	648

RESULT 6	LOCUS	DEFINITION	EST 26-OCT-2001
BB649967	BB649967	625 bp mRNA linear	
	BB649967	RIKEN full-length enriched, 16 days embryo head Mus	
		musculus cDNA clone C130099N17 5', mRNA sequence.	

VERSION	BB649967.1	GI:16484222
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

Eukaryota; Metazoa: Chordata; Cranista; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae: Mus 1 (bases 1 to 625)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

TITLE
JOURNAL
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shuho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-riescgsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Kono, H., Okazaki, Y., Kikunishi, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Waki, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E.,
Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a

FEATURES
source
Location/Qualifiers
1..625

	location/Qualifiers
	1..625
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10990"
	/clone="Cl30099N17"
	/clone_1lb="RIKEN full-length enriched, 16 days embryo head"
	/sex="mixed"
	/tissue_type="head"
	/dev_stage="16 days embryo"
	/lab_host="DH10B"
	/note="Site-1: SalI; Site-2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research with the contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15'
	GAGAGGAGAAGATCCACAGCGCTTTTGTTTTTTTTTTTC 3'. cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15'
	GAGAGGAGATTCGAGTTAAATTAAATTATCCCCCCCCCCC 3'. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FIC I"
BASE COUNT	136 a 198 c 135 g 156 t
ORIGIN	
Query Match	19.8%; Score 429; DB 9; Length 625;
Best Local Similarity	100.0%; Pred. No. 6.2e-40;
Matches 429; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Dy	1 CTTCCCGCGGGGATCGCTTAGAAGTGATATACACGACGATCGATTACCACCGTGA 60
Db	197 CTTCCCGCGGGGATCGCTTAGAAGTGATATACACGACGATCGATTACCACCGTGA 256
Dy	61 GTAGTTCAAGCACTATCTCGCTTTCATGCTATGTTCTGTACTGAAAAATTATGACA 120
Db	257 GTAGTTCAAGCACTATCTCGCTTTCATGCTATGTTCTGTACTGAAAAATTATGACA 316
Dy	121 TCAGATCCAAGGCATTTCTGCCCATTTCTAATCGCTGCATGATCCTGGCGCTTC 180
Db	317 TCAGATCCAAGGCATTTCTGCCCATTTCTAATCGCTGCATCCTGGCGCTTC 376
Dy	181 ATGCGATGTGCTCATTTTATCAAAGCCCAACAGCTGGGTCTTCAGTCGAATGGAG 240
Db	377 ATGCGATGTGCTCATTTTATCAAAGCCCAACAGCTGGGTCTTCAGTCGAATGGAG 436
Dy	241 TCTGCAATGTGCTGGAATAATGAAAATTTCTTCCACAAAACGATATTATTATAC 300
Db	437 TCTGCAATGTGCTGGAATAATGAAAATTTCTTCCACAAAACGATATTATTATAC 496
Dy	301 GAATCTACCATTCCTGGTTTGGGTATGGCATTGGGACAGACTTTGACCTTACCTCTGC 360
Db	497 GAATCTACCATTCCTGGTTTGGGTATGGCATTGGGACAGACTTTGACCTTACCTCTGC 556
Dy	361 CAAGCAATGTTCAATATCAAGGGGTGCCATCTCAACAAGACCGCTCATTTGTACACAA 420
Db	557 CAAGCAATGTTCAATATCAAGGGGTGCCATCTCAACAAGACCGCTCATTTGTACACAA 616
Dy	421 TCCCATGCG 429

[illegible]

DB	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
DB	361	ANGTCAAAAAGCATCAGAAATFATAGTCGTCGTGTAATTTAGAGAAATCGTTTGGAAAT	420					
OY	1193	AAAGTGTCCATCATTTGTCACA---GTGAAAAAGATTAAATGAGACATAATTCAGTTTTC	1248					
DB	421	AAAAATTTTTCATCACTTGCGACACTTGATGAAATATTTTGTAGATATTCACAGATATG	480					
OY	1249	AGGATTAAGAAAGA-----AAGACGCTGCGCATTTGGCGCACCCTGTTAAATTTCCGCGCTCC	1303					
DB	481	AGGATTAAGAAAGATGACATCACTACTTCTTTGTGCACAAATTTATTTATTCACCTCTC	540					
OY	1304	TTGAGACCATCATATATTTTGGT-CAAAATTTTAAAGATGAGAAATTTAGCATTCACCTCA	1362					
DB	541	TAGGTTAAACGTTATATCTTGTGTGAGATTTTAAAAAGCTCAGCATGAGCAATCATTTCA	600					
OY	1363	TTTGGTTTAA 1372						
DB	601	TTNCGCTTAA 610						
RESULT 8								
LOCUS	BB640911							
DEFINITION	BB640911 RIKEN full-length enriched, 7 days neonate cerebellum Mus musculus cDNA clone A730093G08 5', mRNA sequence.							
ACCESSION	BB640911							
VERSION	BB640911.1	GI:16476069						
KEYWORDS	EST.							
SOURCE	house mouse.							
ORGANISM	Mus musculus							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 767)							
AUTHORS	Arakawa, T., Carlinci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, M., Sasak, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.							
TITLE	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)							
JOURNAL	Unpublished (2001)							
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22, Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) vagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carlinci, P., Sugahara, Y., and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamana, I., Atzawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y. Computational analysis of full-length mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)							

QY	1066	TGAGCAAAAGCATTTTACGTCTAAACCTACACGGTTTGGCAATGCATGCATCCGCGCA	1125
Db	399	TGGAGGAAGGATTTCACTCTTAATCTTCCACATTTTGGGAATCAGTCATGTTGGCT	340
QY	1126	TGCGACATGTAAAAAGGCATCACAATATATAGTCTGTTGTTAGTAAATGGTTT	1185
Db	339	TGCGATATGTCGAAAGGCATCAACAATAATAGTCTGTTGTTAGTAAATGGTTT	280
QY	1186	TGGAAAT- AAAGTGTCCATCATTTGTCACA---GTAAAAAGATTAAATGAGCATATTTCA	1241
Db	279	TGGAATTTAAATTTTTTTCATCACTGTCACACTTGCATGTAATATTTTATGAGATATCATCA	220
QY	1242	AGTTTTCAGATTAACAAACA-----AACACGGTCATTTGGGGCACCCTTTAAATTTTCCG	1296
Db	219	AGTATTTGAAATTAAGAAAGATGCACACTACTCTTTTGTGTGCACAAATTTATTTTATCA	160
QY	1297	CCCTCCCTTGAGAGCATTATATTTTGGT- GAATTTTTTAAAGATGAGAATTAAGCAAT	1355
Db	159	CCCTCTTAGGGTAAAGCTGTATATTTTGGTGGAGATTTTAAACCTGAGATGAGCAAT	100
QY	1396	CACCTAATTTGGTTTAAATATATCTGTATATATGTGATTAACGACACTGGA- AATTAAT	1414
Db	99	CATTCAATTCGGTTTAAATTAATTCGTATATACCAATTAATTAATGTGCTGAGAGTAAT	40
QY	1415	TATTTGTCACTCTCATTTATATAACATTTGTTTATACA	1450
Db	39	TATTTTCATATATCATTTTGTAAACATTTGCTTTTCA	4
RESULT 10			
LOCUS	BB651409		
DEFINITION	BB651409 RIKEN full-length enriched, 0 day neonate cerebellum Mus		
ACCESSION	BB651409		
VERSION	BB651409.1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 628)		
	Arakawa,T., Carlinich,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Koude,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasakih,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.		
	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
	Unpublished (2001)		
TITLE	Contact: Yoshihide Hayashizaki		
JOURNAL	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute		
COMMENT	The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/		
	Carnilich,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)		
	waghi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Iwama,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)		

[illegible]

1.

2.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 44.2762 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-3

Perfect score: 2170
Sequence: 1 cttccctcgtcgtggtcgc.....aaaaaaaaaaaaaaaaaaaaa 2170

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128.6	5.9	1814	2	US-08-483-151-1
2	128.6	5.9	1814	5	PCT-US96-06427-1
3	128.2	5.9	1086	1	US-07-914-281-12
4	128.2	5.9	1086	1	US-08-393-246-12
5	128.2	5.9	1086	1	US-08-525-058A-12
6	128.2	5.9	1086	4	US-08-696-731-12
7	128.2	5.9	1086	4	US-09-042-531-12
8	128.2	5.9	2042	3	US-09-063-237-2
9	128.2	5.9	2043	1	US-07-914-281-1
10	128.2	5.9	2043	1	US-08-393-246-1
11	128.2	5.9	2043	1	US-08-273-411-4
12	128.2	5.9	2043	1	US-08-525-058A-1
13	128.2	5.9	2043	2	US-08-696-731-1
14	128.2	5.9	2043	4	US-09-042-531-1
15	128.2	5.9	2043	5	PCT-US91-00899-1
16	126.6	5.8	1316	1	US-07-914-281-10
17	126.6	5.8	1316	1	US-08-393-246-10
18	126.6	5.8	1316	1	US-08-525-058A-10
19	126.6	5.8	1316	2	US-08-696-731-10
20	126.6	5.8	1316	4	US-09-042-531-10
21	124.6	5.7	1654	1	US-07-914-281-13
22	124.6	5.7	1654	1	US-08-393-246-13
23	124.6	5.7	1654	1	US-08-525-058A-13
24	124.6	5.7	1654	2	US-08-696-731-13
25	124.6	5.7	1654	4	US-09-042-531-13
26	120	5.5	1068	5	PCT-US91-00899-13
27	120	5.5	1215	5	PCT-US91-00899-5

28	120	5.5	1488	1	US-07-914-281-9	Sequence 9, Appl
29	120	5.5	1488	1	US-08-393-246-9	Sequence 9, Appl
30	120	5.5	1488	1	US-08-525-058A-9	Sequence 9, Appl
31	120	5.5	1488	2	US-08-696-731-9	Sequence 9, Appl
32	120	5.5	1488	4	US-09-042-531-9	Sequence 9, Appl
33	120	5.5	2175	4	US-08-482-073-9	Sequence 9, Appl
34	120	5.5	2861	4	US-08-482-073-10	Sequence 10, Appl
35	120	5.5	3647	1	US-07-914-281-7	Sequence 7, Appl
36	120	5.5	3647	1	US-08-393-246-7	Sequence 7, Appl
37	120	5.5	3647	2	US-08-525-058A-7	Sequence 7, Appl
38	120	5.5	3647	4	US-08-696-731-7	Sequence 7, Appl
39	120	5.5	3647	5	US-09-042-531-7	Sequence 7, Appl
40	120	5.5	3647	5	PCT-US91-00899-4	Sequence 4, Appl
41	116.8	5.4	2134	2	US-08-483-151-3	Sequence 3, Appl
42	116.8	5.4	2134	5	PCT-US96-06427-3	Sequence 3, Appl
43	67.4	3.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
44	66.2	3.1	6124	4	US-08-213-419B-3	Sequence 3, Appl
45	65.2	3.0	5852	1	US-07-867-106-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-483-151-1
; Sequence 1, Application US/08483151
; Patent No. 5858752

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Holgersson, Jan

TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

SECRET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,151

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Lech, Karen F.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/278001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1814 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-483-151-1

Query Match 5.9%, Score 128.6; DB 2; Length 1814;
Best Local Similarity 52.5%; Pred. No. 7.8e-17;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;
DB 548 GGTGCGTCTGAGTCTAACCGAGCCGTGCTAGCCAGTGTGATGCTGTGCTTCACACC 607

Tue Oct 8 10:18:16 2002

us-09-744-748-3.in1

Page 2

```

OY 443 ATAGACATCAGCTGGATCTGACTAACTTACCTAGCAGGCGCAGCCCTTTCAGA 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 608 ACCGTGAGCTGCAAAACCGGCAATCTCTTACCCCTGGACAGAGCGCACGACGAC 667
OY 503 AATGATTTGATGAATTTAGATGACACCTACACAGCCCGCAAAAGAGGATGAAAC 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 CTGGGTCTGGGCTTCATGATGCGCCGATATGCCATATGCTCCATCGCTTCCGG 727
OY 563 ACTTGTCAACCTGACTTAACTTATCCGCTGATTCAGATATCCAAAGCTTATGCT 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 GCATCTTCAACCTGGGTGCTGAGCTGAGGCTGATTCAGATATCTTTGATCCCTAGGCTC 787
OY 623 TCTTGACGCTGACCAAAATCCCTTGTGTTGATGAGTGCACAGCAAGAGAGTGTGT 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 GCTTGAGAGCTCTCTCTGGGCC--ACATCCCACTACCGGCAAAAGAGATGGCTG 844
OY 683 GCTGGTGTGATGATGAGAACCTGAGCATGCCAGGCTGATATTAAGAGAGCTCA 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 845 CTTGGTGTGATGAGCAATTTCCAGAGAGGCGAGAGGCTGCAAGCTGACCGACCTGG 904
OY 743 GCAAGATATGAAATCCACACCTATGCGCCAAAGCTTGGAGAAATACGTGAACGATAAA 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 905 CCCCTCATCTGAGGTGATGTGTCGTCGCCAGCGAGCGCCCTATGCGCTAATTT 964
OY 803 ATCTGATCCCACTATCTACTTGTAAATTTATCTTTCATTTGAAAAGTCAATTCACA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 965 GTCTGCTGCCCACTTGTGGCCGCTACCGCTTCTTACCTGACCTTGGAGACGTCAACAGCATC 1024
OY 863 AAGATTCATCAGACAGAAAAGCTTA--CAATGCAATTTTGGCTGCTGACGCTGTG 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1025 GGGACTACATCACTAGAAAGTTCTGGGCAATGCTTGGCGGTGTGTGCTGACCTGG 1084
OY 920 TCTGGGTCTCATCTAGGAAAACCTATGAGATTAATTTCCAGCTGATTCATTCATG 979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1085 CGTGGAGCTCTCTGGGCACTGACGAGCTTTGTGTGCAACAGATGCTTTGTACAG 1144
OY 980 TGAAGATTTTAACTCTCCAGTAGATTAGC 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1145 TGAAGACTTCACTCTGCTGCGGTGAACTGGC 1175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
PCT-US96-06427-1
: Sequence 1, Application PC/TUS9606427
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation
: TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USBS THEREOF
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/06427
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/483,151
: FILING DATE: 07-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lech, Karen F.
: REGISTRATION NUMBER: 35,238
: REFERENCE/DOCKET NUMBER: 00786/278W01
: TELECOMMUNICATION INFORMATION:

```

```

: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SFO ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1814 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
PCT-US96-06427-1

```

```

Query Match 5.9%; Score 128.6; DB 5; Length 1814;
Best local similarity 52.5%; Pred. No. 7.6e-17;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

```

```

OY 383 GCTCCATCTCACAAGACCGCTCATTTGTACAAATCCATGCGGTCTGATTAC 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 548 GCTGCCGTGTGATGCTTAACCGGAGCTGTAGCCAGTGTGATGCTGTGCTTCCAC 607
OY 443 ATAGACATCAGCTGGATCTGACTAACTTACCTAGCAGGCGCAGCCCTTTCAGA 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 608 ACCGTGAGCTGCAAAACCGGCAATCTCTTACCCCTGGACAGAGCGCACGACGAC 667
OY 503 AATGATTTGATGAATTTAGATGACACCTGAGCATGCCAGGCTGATATTAAGAGAGCTCA 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 CTGGGTCTGGGCTTCATGATGCGCCGATATGCCATATGCTCCATCGCTTCCGG 727
OY 563 ACTTGTCAACCTGACTTAACTTATCCGCTGATTCAGATATCCAAAGCTTATGCT 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 GCATCTTCAACCTGGGTGCTGAGCTGAGGCTGATTCAGATATCTTTGATCCCTAGGCTC 787
OY 623 TCTTGACGCTGACCAAAATCCCTTGTGTTGATGAGTGCACAGCAAGAGAGTGTGT 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 GCTTGAGAGCTCTCTCTGGGCC--ACATCCCACTACCGGCAAAAGAGATGGCTG 844
OY 683 GCTGGTGTGATGATGAGAACCTGAGCATGCCAGGCTGATATTAAGAGAGCTCA 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 845 CTTGGTGTGATGAGCAATTTCCAGAGAGGCGAGAGGCTGCAAGCTGACCGACCTGG 904
OY 743 GCAAGATATGAAATCCACACCTATGCGCCAAAGCTTGGAGAAATACGTGAACGATAAA 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 905 CCCCTCATCTGAGGTGATGTGTCGTCGCCAGCGAGCGCCCTATGCGCTAATTT 964
OY 803 ATCTGATCCCACTATCTACTTGTAAATTTATCTTTCATTTGAAAAGTCAATTCACA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 965 GTCTGCTGCCCACTTGTGGCCGCTACCGCTTCTTACCTGACCTTGGAGACGTCAACAGCATC 1024
OY 863 AAGATTCATCAGACAGAAAAGCTTA--CAATGCAATTTTGGCTGCTGACGCTGTG 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1025 GGGACTACATCACTAGAAAGTTCTGGGCAATGCTTGGCGGTGTGTGCTGACCTGG 1084
OY 920 TCTGGGTCTCATCTAGGAAAACCTATGAGATTAATTTCCAGCTGATTCATTCATG 979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1085 CGTGGAGCTCTCTGGGCACTGACGAGCTTTGTGTGCAACAGATGCTTTGTACAG 1144
OY 980 TGAAGATTTTAACTCTCCAGTAGATTAGC 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1145 TGAAGACTTCACTCTGCTGCGGTGAACTGGC 1175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
CORRESPONDENCE ADDRESS: 14
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696,731
FILING DATE: 14-AUG-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
FILING DATE: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-696-731-12
Query Match 5.9%; Score 128.2; DB 2; length 1086;
Best Local Similarity 53.0%; Pred. No. 8.5e-17;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

733 ACGAGCTCAGCAAGATATGAAATCCACACCTATGCGCCAGCATTCGAGAAATACGTG 792
631 CAGAGCCTGAGGCTCATCTCAAGAGGTGAGGTATGAGGAGGCTCCACCA---AGCCCTG 687
793 ACGATTAATAATGATTCCTCCACCATATCTACTTGTAAATTTATCTTCAATTTGAAAC 852
688 CCCAAGGGGACCATGATGAGAGCGTGTCCGGGTACAAAGTTTACTGCGCTCGAAGAC 747
853 TCAATTCACAAAGATTAATACATACAGAAAGCTCT---ACATCATTTTGGCTGTTCA 909
748 TCCCTGACCCCGACTACATACACAGAAAGCTTGAGAGAAAGCCCTGAGGCTGGCC 807
910 GTACCTGTGTCTCCTGGTCCATCTAGGAAAGTATGAAATATATTCACGCTGATTA 969
808 GTCCCGGTGTGTGTGGCCCGACAGAGAACTAGAGAGGTCTCTGCCACCCGACGCC 867
970 TTCAATTCATGAGAAATTTAACTCTCCAGTGAAGTAAATATCTGAAGAAAGTT 1029
868 TTCAATTCATGAGAAATTTAACTCTCCAGTGAAGTAAATATCTGAAGAAAGTT 927
1030 GACAAATAATATGTTGATCTTACTTACTTAACTCGAAGAG 1074
928 GACAAAGACCAAGCCCGCTACCTGAGCTTCTGCTGGGGAG 972
RESULT 7
US-09-042-531-12
Sequence 12, Application US/09042531
Patent No. 6268193
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
CORRESPONDENCE ADDRESS: 14
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
FILING DATE: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid

Tue Oct 8 10:18:16 2002

STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-12

Query Match
Best Local Similarity 5.9%; Score 128.2; DB 4; Length 1086;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

Db 271 TGGCATTCTCAGACAGACCGCTGATTTGACCAAAATCCCATGCGTCTGATTCACCAT 444
OY 445 AGAGACATCAGCTGGGGA--TTCGACTTAATCTTACAGAGAGAGAGAGAGAGAGAGAG 330
Db 331 TGGGATATCATGTCGCAACCCCTTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
OY 502 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
Db 391 CCGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
OY 562 CACTGTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
Db 451 AGATCTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
OY 622 TTTTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
Db 511 TGGTGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
OY 673 AAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
Db 571 GAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
OY 733 AAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
Db 631 CAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
OY 793 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Db 688 CCCAAGGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
OY 853 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
Db 748 TCTTGGACCCCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
OY 910 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Db 808 GTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
OY 970 TTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
Db 868 TTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
OY 1030 GACAAAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
Db 928 GACAAAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

RESULT 8
US-09-063-237-2
Sequence No. 6124267
GENERAL INFORMATION:
APPLICANT: McEvoy, Rodger P.
APPLICANT: Cummins, Richard D.
TITLE OF INVENTION: O-Glycan Inhibitors of Selectin Mediated
CORRESPONDENCE ADDRESS: Inflammation Derived from PSL-1
ADDRESS: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta

us-09-744-748-3.rn1

STATE: Georgia
COUNTRY: US
ZIP: 30306-3450
COMPUTER READABILITY: FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA: Version #1.25
FILING DATE: 08/6/94
CLASSIFICATION: US-09/063.237
PRIORITY APPLICATION: 08/6/94, 802
FILING DATE: 08/6/94, 802
ATTORNEY/AGENT INFORMATION:
NAME: Patrea L. Pabst
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: 08/6/94, 802
TELEPHONE: (404) 873-8795
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-063-237-2

Query Match
Best Local Similarity 5.9%; Score 128.2; DB 3; Length 2042;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

OY 385 TGGCATTCTCAGACAGACCGCTGATTTGACCAAAATCCCATGCGTCTGATTCACCAT 444
Db 343 TGGCATTCTCAGACAGACCGCTGATTTGACCAAAATCCCATGCGTCTGATTCACCAT 444
OY 445 AGAGACATCAGCTGGGGA--TTCGACTTAATCTTACAGAGAGAGAGAGAGAGAGAGAG 402
Db 403 TGGGATATCATGTCGCAACCCCTTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
OY 502 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
Db 502 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
OY 463 CCGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
Db 463 CCGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
OY 562 CACTGTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
Db 523 AGATCTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 522
OY 523 AGATCTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
Db 622 TTTTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 582
OY 583 TGGTGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
Db 673 AAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
OY 643 GAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
Db 703 CAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
OY 733 AAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
Db 793 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
OY 760 CCCAAGGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
Db 853 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
OY 820 TCTTGGACCCCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
Db 910 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879

Db 880 GTGCGCGTGTGCTGGGCCCCAGACAGACTACAGAGAGGTTCTGCGACCCGAGGCC 939
QY 970 TTGATTCATGTGAGAGATTTTAACTCTCCAGTAGTACAGAAATATGAGGAAGTT 1029
Db 940 TTGATTCAGGTGAGACACTTCCAGAGCCCAAGAGACTGCGCGGATCTGAGAGAGCTG 999
QY 1030 GACAAACAAATAGTTAGTCTAGTACTTAACTGAGGAAG 1074
Db 1000 GACAAAGACACGCGCCGCTACCTGAGCTACTTCTGCTGGCGGAG 1044

RESULT 9

US-07-914-281-1
; Sequence 1, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOME, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,281
; FILING DATE: 19920720
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAR UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2043 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ANTI-SENSE: NO
; US-07-914-281-1

Query Match 5.9%; Score 128.2; DB 1; Length 2043;

Best Local Similarity 53.0%; Pred. No. 9.5e-17;

Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 385 TGCATCTGACAAAGCCGCTATGTACAAACAAATCCAGCGTCTGATTCACCAT 444
Db 343 TGCACATCAGTCCGACGCAAGGTATCCACAGCAGACAGCAGTATCTGACACAC 402
QY 445 AGAGACATGAGTGGGA---TCTGACTAATCTACTGACAGGCCAGGCCCTTTGAG 501
Db 403 TGGGATATCATGTCCAAACCTTAAGTACGCTCCACCTTCCCGGCGGAGGGGCGAG 462
QY 502 AATGATTTGATGATTTAGAGTACACCCAGTACACCCCAAGAGAGTGGCATTTGAA 561
Db 463 CGGTGATCTGTTCAACTTGGAGCCACCCCTTAAGTGCACGACCTGGAAGGCTTGAGC 522

QY 562 CACTGTTCACCTGACTCTAATCTACGCGGTGATTCAGATATCCAGTGCCTTATGCG 621
Db 523 AGATATCTCAATCTCAACATGTCCTACCGGAGGACTCCGACATCTTCACGCGCTACGG 582
QY 622 TTCTTACGCTGAGCACAATATCCCTTTG-----TGTTCAGTGCACAGCAAGAG 672
Db 583 TGGCTGAGGCGGTGTCCGCGCCAGCCTGCCACCCAGCTCAACCTCTCGGCGCAAGACC 642
QY 673 AAGTTGTGCTGGGTGTGAGTAACTGGAACCTGAGCATGCGGAGGTCAGTATTAC 732
Db 643 GAGCTGTGCTGGGTGTGAGTAACTGGAACCTGAGCATGCGGAGGTCAGTATTAC 702
QY 733 AACGACTCAGCAAGATATTGAAATTCACACCTATGCGCAGCATTCGAGAGATGAG 792
Db 703 CAGAGCTCAGAGCTCATCTCAAGGTGAGAGTGTACGAGAGCTGCCACA---AGCCCTTG 759
QY 793 AACGATAAATATGATTTCCACCATATCTACTTGAATTTATCTTTCATTTGAAAC 852
Db 760 CCCAAGGGGACCATGATGAGAGAGCTGCCGCTACAGTTCTACTGCTTGCAGAAC 819
QY 853 TCATTCACAAAGATATACATCAGAGAAAGCTCT---ACATGCAATTTTGGCTGTCA 909
Db 820 TCTTTCACCCCGCAGCTACATCACCAGAAAGCTGTGAGAGAGCCCTGAGGCTTGCGC 879
QY 910 GTACCTGTGTCTGCTGCTCATCTAGGAAACTATGAGATTTATTCAGCTGATTC 969
Db 880 GTGCGCGTGTGCTGGGCCCCAGCAGAGAACTAGAGAGTCTCGCACCCGAGCGCC 939
QY 970 TTGATTCATGTGAGAGATTTTAACTCTCCAGTAGTACAGAAATATGAGGAAGTT 1029
Db 940 TTGATTCAGGTGAGACACTTCCAGAGCCCAAGAGACTGCGCGGATCTGAGAGAGCTG 999
QY 1030 GACAAACAAATAGTTAGTCTAGTACTTAACTGAGGAAG 1074
Db 1000 GACAAAGACACGCGCCGCTACCTGAGCTACTTCTGCTGGCGGAG 1044

RESULT 10

US-08-393-246-1

; Sequence 1, Application US/08393246

; Patent No. 5595900

; GENERAL INFORMATION:
; APPLICANT: LOME, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451

Tue Oct 8 1...:..6 2002

us-09-744-748-3.rn1

REFERENCE/DOCKET NUMBER: 2363-060-55
TELEPHONE: (703)521-4500
TELETYPE: (703)486-2347
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: unknown
ANTI-SENSE: NO
US-08-393-246-1

Query Match
Best Local
Matches 374; Conservative 53.08; Score 128.2; DB 1; Length 2043;
Mismatches 313; Indels 18; Gaps 4;

OY 385 TGGCATCTGCAACAGACGCTCATCTTACACAAATCCCATGCGCTGCTGATTCACCAT 444
DB 343 TGGCATCTGCAACAGACGCTCATCTTACACAAATCCCATGCGCTGCTGATTCACCAT 444
OY 445 AGAGACATGACGCGCA---TCTGACTAATCTTACTGACAGAGCCAGCCAGCCCTTTTCA 402
DB 403 TGGCATCTGCAACAGACGCTCATCTTACACAAATCCCATGCGCTGCTGATTCACCAT 402
OY 502 AATGATTTGATGATTTTAAAGTACCCACTGACAGCCCGCCAGCCAGCCAGCCAGCCAG 501
DB 463 CGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 462
OY 562 CACTGTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
DB 523 AGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522
OY 622 TTTCTGAGGTGACGCAAAATCCCTTTC-----TGTGAGTGGCAAGCAAGAG 621
DB 583 TGGCATCTGCAACAGACGCTCATCTTACACAAATCCCATGCGCTGCTGATTCACCAT 582
OY 673 AAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 672
DB 643 GAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 642
OY 733 AACGAGCTGACGCAAGATTTGAATTCACACACTGAGAGAGAGAGAGAGAGAGAGAGAG 732
DB 703 CAGAGCTGACGCAAGATTTGAATTCACACACTGAGAGAGAGAGAGAGAGAGAGAGAG 702
OY 793 AACGATTTGATGATTTTAAAGTACCCACTGACAGCCCGCCAGCCAGCCAGCCAGCCAG 792
DB 760 CCCAAGGAGACGATGAG 759
OY 853 TCAATTTACAGATTTACATTCACAGAAAGCTCT---ACATGACTTTTGGCTGCTTCA 852
DB 820 TCTTGGACCGCCGACATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
OY 910 GTACCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
DB 880 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
OY 970 TTTGATTCATGAGATTTTAAAGTACCCACTGACAGCCCGCCAGCCAGCCAGCCAGCCAG 969
DB 940 TTTGATTCATGAGATTTTAAAGTACCCACTGACAGCCCGCCAGCCAGCCAGCCAGCCAG 939
OY 1030 GACAAAAAATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1029
DB 1000 GACAAAGACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 999

RESULT 11
US-08-273-411-4
Sequence 4, Application US/08273411
Grant No. 5625124
GENERAL INFORMATION:

APPLICANT: Falk, Per
TITLE OF INVENTION: Animal Model for Gastro-Intestinal
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS: 10
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/273,411
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Patent, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: W0106
TELEPHONE: (404) 815-6508
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE: NO
NAME/KEY: misc.feature
LOCATION: 1..361
OTHER INFORMATION: /note= "Nucleotides 73 through 1158 encode the GDP-L-fuco
PUBLICATION INFORMATION:
AUTHORS: Kukowska-Latallo, et al.
JOURNAL: Genes & Development
VOLUME: 4
PAGES: 1288-1303
DATE: 1990
RELEVANT RBD
US-08-273-411-4
RELEVANT RBD
US-08-273-411-4
Query Match
Best Local
Matches 374; Conservative 53.08; Score 128.2; DB 1; Length 2043;
Mismatches 313; Indels 18; Gaps 4;

```

QY 673 AAGTGTGTGCTGGTGTGTGAGTAACGTGAACCTGACATGCCAGGCTCAAGTATTAC 732
    ||||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 643 GAGCTGTGGCTGGGCGGTGTCAACTGGAAGCCGAGCTCAGCCAGGCTGAGCTACTAC 702
QY 733 AACGAGCTCAGCAAGAGTATTGAATCCACACCTATGCCAAGCATTTGGAGAAATACG 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 703 CAGAGCTGTGAGCTCATCTCAAGGTGAGCTGTACGGAGCTCCACA---AGCCCCCTG 759
QY 793 AACGATAAATATCTGATCTCCACCATATCTACTTGAATTTATCTTCAATTGGAAC 852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 760 CCCAAGGGACCATGATGAGAGAGCTGTCCCGGTACAGTTCTACTGGCTTCGAGAAC 819
QY 853 TCAATTCACAAAGATTACATCAACAGAAAGCTCT---ACAATGCAATTTTGGCTGATCA 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 TCCCTGCACCCGACTACATCAACGAGAAAGCTGTGAGAGAACGCCCTGAGGCTGGGCC 879
QY 910 GTACCTGTGTCTCTGGGTCACTGAGGAAACATGAGAAATATATTTCCAGCTGATTCA 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 GTGCCCCGTGTCTGTGGCCCCCAGCAGAAAGCACTACGAGAGGTCTCCGCCACCCGACGCC 939
QY 970 TTCATTCAATGAGAGATTATTAACCTCCACAGTAGTACCAAAATATCTGAAGAGTT 1029
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 940 TTCATTCACGTGAGAGACTTCCAGAGCCCAAGAGACTGGCCCGGTACTCGAGAGCTG 999
QY 1030 GACAAACCAATATAGTTGATACCTTAGTTACTTTAATCTGAGAGAAG 1074
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1000 GACAAGAGACACGCGCGGTACTGAGACTTTCGCTGGCGGGAG 1044

```

RESULT 12

US-08-525-058A-1
Sequence 1, Application US/08525058A
Patent No. 5770420

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P. C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525, 058A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ANTI-SENSE: NO

US-08-525-058A-1

Query Match 5.9%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred No. 9,5e-17;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

```

QY 385 TGCATCTTCACAACAGACCGCTATTTGTACAAACAAATCCATGCGGTCTGATTCCAT 444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 343 TCCACATCATCTGCGAGCCGCAAGGTATCCACAGGAGACAGGTATCGTGCACGAC 402
QY 445 AGAGCATATAGCTGGGA---TCTGACTTAATCTTACCTGACAGGCCAGCCCTTTGAG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 TGGATATCATGTCCAAACCTTAAGTCACGCTCCACCTTCCCGGAGCCGAGGGGAG 462
QY 502 AATGATTTGGATGATGATTTAGATGATACACCTACACACCCCAAAAGAGTGGCAATTGA 561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 CCCTGATCTGTTCACTTGTGAGAGCCACCCCTTAATGTCAGACCTGGAAACCCCTGGAC 522
QY 562 CACTGTTCACCTGACTCTTAATCTATCCGCTGATTCAGATATCCAAAGTGCCTATGAC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 523 AATATCTTCAATCTCAACATGTCATACGAGAGACTCCGACATCTTCACGCCCTAGAGC 582
QY 622 TTCTTGACGCTGACCAATATCCCTTTG-----TGTGAAAGTCCAAAGCAAGAG 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 583 TGGCTGAGACCTGTGTGTCGCGCCAGCCTGCCACCCACCCCTCAACCTCTCGGCCAAGACC 642
QY 673 AAGTGTGTGCTGGGTTGTGAGTAACGTGAACCCCTGACATGCCAGGCTCAAGTATTAC 732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 643 GAGCTGTGGCTGGGCGGTGTCAACTGGAAGCCGAGCTCAGCCAGGCTGAGCTACTAC 702
QY 733 AACGAGCTCAGCAAGAGTATTGAATCCACACCTATGAGCAAGCATTTGGAGAAATACG 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 703 CAGAGCTGTGAGAGCTCATCTCAAGGTGAGAGCTGTACGGAGAGCTCCACA---AGCCCCG 759
QY 793 AACGATAAATATCTGATCTCCACCATATCTACTTGAATTTATCTTCAATTGGAAC 852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 760 CCCAAGGGACCATGATGAGAGAGCTGTCCCGGTACAAATTTCTACTGGCTTCGAGAAC 819
QY 853 TCAATTCACAAAGATTACATCAACAGAAAGCTCT---ACAATGCAATTTTGGCTGATCA 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 TCCCTGCACCCGACTACATCAACGAGAGCTGTGAGAGAACGCCCTGAGGCTGGGCC 879
QY 910 GTACCTGTGTCTCTGGGTCACTGAGGAAACATGAGAAATATATTTCCAGCTGATTCA 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 GTGCCCCGTGTCTGTGGCCCCCAGCAGAAAGCACTACAGAGGTTCCTGCCACCCGACGCC 939
QY 970 TTCATTCAATGAGAGATTATTAACCTCCACAGTAGTACCAAAATATCTGAAGAGATT 1029
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 940 TTCATTCACGTGAGAGACTTCCAGAGCCCAAGAGACTGGCCCGGTACTCGAGAGCTG 999
QY 1030 GACAAACCAATATAGTTGATACCTTAGTTACTTTAATCTGAGAGAAG 1074
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1000 GACAAGAGACACGCGCGGTACTGAGACTTTCGCTGGCGGGAG 1044

```

RESULT 13

US-08-696-731-1
Sequence 1, Application US/08696731
Patent No. 595347

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P. C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.


```

Db 343 TGCACATCATCTGCGACCGGACAGGTGTACCCACAGCAGCAGCGTCATGTGCACACC 402
QY 445 AGACATCATCTGCGACCGGACAGGTGTACCCACAGCAGCAGCGTCATGTGCACACC 501
Db 403 TGGATATCATGTGCAACCTTAAGTACAGCTCCACCTTCCGAGGCGCCAGAGGGCAG 462
QY 502 AATGATTTGGATGAATTTAGAGTCAACCCACTCACACCCCAAAAAGATGGCATTTGA 561
Db 463 CGCTGATCTGCTTCAACCTTGAGCACAACCCCTTAAGTCCAGCAGCAGCGTGGAC 522
QY 562 CACTTGTCAACCTGACTTACTTATGCGCGTGTATTCAGATTCAGATTCAGTTCAGC 621
Db 523 AGATACCTTCAATCTCACAGTCTCTACCGAGCAGCTCCGATCTTACGCGCTTACGCG 582
QY 622 TTTTGAAGGAGCAAAATCCCTTG-----TGTGTAAGTGGCAAGCAAGAG 672
Db 583 TGGCTGAGACCGGTGTGCGGCGACCGTCCACCCACCGCTCAACCTGTGGCCAAAGACC 642
QY 673 AAGTTGGTGTGCTGGTGTGAGTACTGNAACCTGAGCATGCGCAGGATCAAGTATTAC 732
Db 643 GAGCTGGTGGCGTGGGCGGTGTCCAACTGGAAGCGGAGCTCAGCAGGATGGCTACTAC 702
QY 733 AAGAGCTCAGCAAGATATTGAATCCACACTATGGCCAAAGCATTCGGAGATACGTG 792
Db 703 CAGAGCTGAGGCTCATCTCAAGGTGAGTGTACGAGCTGCCACA--AGCCCTG 759
QY 793 AAGATTAATAATGTATCCACCATCTACTGTAAATTTTATCTTATTTGAAAC 852
Db 760 CCAAGAGGACCATGATGAGAGAGCTGTCCCGGTACAAATTTTACTGAGCTTGAAGAC 819
QY 853 TCAATTCACAAAGATTCATACAGAAAGCTCT--ACAATGATTTTGGCTGTCTCA 909
Db 820 TCTTGACACCCGACATACACGAGAGAGAGTGTGAGGAAAGCGCCGTGGAGCTGGGCC 879
QY 910 GTACTGTGTCTGGTCCATCTAGGAAACATATGAGATTTATTCAGCTGATTTCA 969
Db 880 GTGCGCTGTGTGCTGGGCGCCAGCAGAGCACTACGAGAGGTCTCCGCAACCCGAGGCC 939
QY 970 TTTATGATGTGGAAGATTTTAACTTCCAGTGTAGCAAAATTCGAGGAAGTT 1029
Db 940 TTATTCACAGTGAAGCTTCCAGAGCCCAAGAGACTGGCCGCTGACTGCAAGAGCTG 999
QY 1030 GACAAAAACAATAAGTTGTAACCTTAGTTACTTTAACTGAGGAAG 1074
Db 1000 GACAAAGACCAAGCCCGCTACCTGAGCTTGTGCGGGGAG 1044

```

RESULT 15
PCT-US91-00899-1
Sequence 1, Application PC/TUS9100899
GENERAL INFORMATION:
APPLICANT: Lowe, John B.
TITLE OF INVENTION: Method and Products For the Synthesis of
TITLE OF INVENTION: Oligosaccharide Structures on Glycolipids,
TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/00899
FILING DATE: 19910214

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye Ph.D., Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
TELEPHONE: (703)521-5940
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SRO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Blood
CELL LINE: A431
PCT-US91-00899-1
Query Match 5.9%; Score 128.2; DB 5; Length 2043;
Best Local Similarity 53.0%; Pred. No. 9.5e-17;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
QY 385 TGCATCTCACACAGACCGCTCATTTTACACAAATCCATGCGGCTGATTCACCAT 444
Db 343 TGCACATCATCTGCGACCGGACAGGTGTACCCACAGCAGCAGCGTCATGTGCACACC 402
QY 445 AGACATCATCTGCGACCGGACAGGTGTACCCACAGCAGCAGCGTCATGTGCACACC 501
Db 403 TGGATATCATGTGCAACCTTAAGTACAGCTCCACCTTCCGAGGCGCCAGAGGGCAG 462
QY 502 AATGATTTGGATGAATTTAGAGTCAACCCACTCACACCCCAAAAAGATGGCATTTGA 561
Db 463 CGCTGATCTGCTTCAACCTTGAGCACAACCCCTTAAGTCCAGCAGCAGCGTGGAC 522
QY 562 CACTTGTCAACCTGACTTACTTATGCGCGTGTATTCAGATTCAGATTCAGTTCAGC 621
Db 523 AGATACCTTCAATCTCACAGTCTCTACCGAGCAGCTCCGATCTTACGCGCTTACGCG 582
QY 622 TTTTGAAGGAGCAAAATCCCTTG-----TGTGTAAGTGGCAAGCAAGAG 672
Db 583 TGGCTGAGACCGGTGTGCGGCGACCGTCCACCCACCGCTCAACCTGTGGCCAAAGACC 642
QY 673 AAGTTGGTGTGCTGGTGTGAGTACTGNAACCTGAGCATGCGCAGGATCAAGTATTAC 732
Db 643 GAGCTGGTGGCGTGGGCGGTGTCCAACTGGAAGCGGAGCTCAGCAGGATGGCTACTAC 702
QY 733 AAGAGCTCAGCAAGATATTGAATCCACACTATGGCCAAAGCATTCGGAGATACGTG 792
Db 703 CAGAGCTGAGGCTCATCTCAAGGTGAGTGTACGAGCTGCCACA--AGCCCTG 759
QY 793 AAGATTAATAATGTATCCACCATCTACTGTAAATTTTATCTTATTTGAAAC 852
Db 760 CCAAGAGGACCATGATGAGAGAGCTGTCCCGGTACAAATTTTACTGAGCTTGAAGAC 819
QY 853 TCAATTCACAAAGATTCATACAGAAAGCTCT--ACAATGATTTTGGCTGTCTCA 909
Db 820 TCTTGACACCCGACATACACGAGAGAGTGTGAGGAAAGCGCCGTGGAGCTGGGCC 879
QY 910 GTACTGTGTCTGGTCCATCTAGGAAACATATGAGATTTATTCAGCTGATTTCA 969
Db 880 GTGCGCTGTGTGCTGGGCGCCAGCAGAGCACTACGAGAGGTCTCCGCAACCCGAGGCC 939
QY 970 TTTATGATGTGGAAGATTTTAACTTCCAGTGTAGCAAAATTCGAGGAAGTT 1029
Db 940 TTATTCACAGTGAAGCTTCCAGAGCCCAAGAGACTGGCCGCTGACTGCAAGAGCTG 999
QY 1030 GACAAAAACAATAAGTTGTAACCTTAGTTACTTTAACTGAGGAAG 1074

```

Tue Oct 8 10:18:16 2002

us-09-744-748-3.rni

Page 12

Db 1000 GACAGGACCCAGCCCGCTACCTGAGCTACTTTCGCTGGCGGAG 1044

Search completed: October 6, 2002, 07:20:16
Job time : 54.2762 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:39:45 ; Search time 1351.36 Seconds
(without alignments)
16699.717 Million cell updates/sec

Title: us-09-744-748-3_COPY_115_1194

Perfect score: 1080
Sequence: 1 atgacatcaacatccaaag.....agaatcgttttggaattaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_da: *
2: gb_hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length DB	ID	Description
------------	-------------	-----------	----	-------------

1	1080	100.0	2139	10	AB015426	AB015426 Mus muscu
2	1008	93.3	1128	10	AF345993	AF345993 Rattus no
3	1004.8	93.0	2156	10	AB049819	AB049819 Rattus no
4	974.4	90.2	1705	10	AF230460	AF230460 Cricetulu
5	936	86.7	2501	9	HS238701	HS238701 Homo sapi
6	932.8	86.4	80247	9	AL512406	AL512406 Homo sapi
7	931.2	86.2	3019	9	AB023021	AB023021 Homo sapi
8	740	68.5	71396	2	AC100674	AC100674 Mus muscu
9	727.6	67.4	1707	5	AB035906	AB035906 Gallus ga
10	675.2	62.5	1080	5	AB035905	AB035905 Xenopus 1
11	268.2	24.8	1567	5	AB023627	AB023627 Dario rer
12	228.2	21.1	1832	5	AB023628	AB023628 Dario rer
13	220.6	20.4	1264	5	AF288369	AF288369 Gallus ga
14	206.6	19.1	1092	5	AB035907	AB035907 Xenopus 1
15	175.2	16.2	2657	5	GGU73678	GGU73678 Gallus gall
16	157	14.5	1429	10	CGU78737	CGU78737 Cricetulu
17	157	14.5	1462	10	AF090450	AF090450 Cricetulu
18	155.4	14.4	1718	10	AF090449	AF090449 Cricetulu
19	145.6	13.5	71396	2	AC100674	AC100674 Mus muscu
20	135.6	12.6	1044	5	AB035908	AB035908 Gallus ga
21	134.8	12.5	2293	4	BTAL34FPG	BTAL34FPG Bos tauru
22	134.8	12.5	12607	4	BTAL32772	BTAL32772 Bos tauru
23	134.2	12.4	1158	10	AB039126	AB039126 Mus muscu
24	133.4	12.4	1113	9	AF345884	AF345884 Saimiti s
25	133	12.3	1119	9	AF345881	AF345881 Macaca mu
26	132.6	12.3	1170	10	AB039127	AB039127 Mus muscu
27	132.6	12.3	1170	10	AB039131	AB039131 Mus muscu
28	132.6	12.3	1615	10	MMU33457	MMU33457 Mus muscu
29	131.4	12.2	1170	10	AB039133	AB039133 Mus spicl
30	131	12.1	1170	10	AB039124	AB039124 Mus muscu
31	131	12.1	1170	10	AB039125	AB039125 Mus muscu
32	131	12.1	1170	10	AB039128	AB039128 Mus muscu
33	131	12.1	1170	10	AB039129	AB039129 Mus muscu
34	131	12.1	1170	10	AB039130	AB039130 Mus muscu
35	131	12.1	1170	10	AB039132	AB039132 Mus muscu
36	130.8	12.1	1302	10	AB049938	AB049938 Rattus no
37	130.8	12.1	169446	2	AC097778	AC097778 Rattus no
38	130	12.0	1613	4	BTAL32776	BTAL32776 Bos tauru
39	130	12.0	1619	4	BTAL32775	BTAL32775 Bos tauru
40	130	12.0	1644	4	BTAL32774	BTAL32774 Bos tauru
41	130	12.0	1873	4	BTAL32773	BTAL32773 Bos tauru
42	129.8	12.0	1162	9	FTY14033	FTY14033 Pan troglod
43	129.8	12.0	80893	9	AC024592	AC024592 Homo sapi
44	129.8	12.0	187488	2	AC010454	AC010454 Homo sapi
45	128.8	11.9	169893	2	AC021165	AC021165 Homo sapi

ALIGNMENTS

RESULT 1
AB015426
LOCUS
DEFINITION Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.
ACCESSION AB015426
VERSION AB015426.1 GI:3702718
KEYWORDS Fut9: alpha1,3-fucosyltransferase IX.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Kudo,T., Ikehara,Y., Togyachi,A., Kaneko,M., Hiraga,T., Sasaki,K.
TITLE Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mfuc-tix, that synthesizes the Lewis x (CD15) epitope in brain and kidney
JOURNAL J. Biol. Chem. 273 (41), 26729-26738 (1998)
MEDLINE 98434588
REFERENCE 2 (bases 1 to 2139)
AUTHORS Kudo,T. and Narimatsu,H.
TITLE Direct Submission

JOURNAL Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science,
Soka University, Division of Cell Biology; 1-236 Tangi-cho,
Hachioji, Tokyo 192-8577, Japan (E-mail:tkudo@t.soka.ac.jp,
Tel:+81-426-91-2495, Fax:+81-426-91-9315)

FEATURES

Location/Qualifiers

1..2139
/organism="Mus musculus"
/db_xref="taxon:10090"
115..1194
/gene="Fut9"
115..1194
/gene="Fut9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="BAA33522.1"
/translation="MTSTSKGILRPFLIVCIILGCFNACLLIYIKPTNSWVSPMESA
SSYLKMKNFPSKTDYFNETTILVWVFPQGFDTLSCAMFNIOGCHLTDLSLNK
SHAVLIHRDLSMDLNLPOQARPPPOKIMMNLSPHTPOKSGIELNLTLYR
DSIDIOVYGLIVSTNPFEVEPSEKELVWVSNMNBHARVRYNELSKSEIHTY
GOAEYVNDKRLIPTISCKEYLSFENSIRKDYTELKNAFLASGYPVLAGPSREN
YENTIPADSEFIHVEDFNSDELAKYIKVDKRNKLYLSYFNMRKDTVNLPRMESH
CLADHYKRIHQEKYSYGNLEKWFNA"
2139

CDS

polyA_site

BASE COUNT 684 a 395 c 368 g 692 t
ORIGIN

Query Match 100.0%; Score 1080; DB 10; Length 2139;
Best Local Similarity 100.0%; Pred. No. 3.1e-278;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACATCAACATCCAAAGCATCTTCCCATTTCTATAGTGTGATCATCCTGGGC 60
DB 115 ATGACATCAACATCCAAAGCATCTTCCCATTTCTATAGTGTGATCATCCTGGGC 174
QY 61 TGCCTCAGCATGCTCTGCTCATTTATATCAAGCCCAACAGAGCTGGTCTCAATCA 120
DB 175 TGCCTCAGCATGCTCTGCTCATTTATATCAAGCCCAACAGAGCTGGTCTCAATCA 234
QY 121 ATGAGAGTCGCAAGTCTGCTGCAAAATGAAAATTTCTGTCACAAAAAAGTATTAT 180
DB 235 ATGAGAGTCGCAAGTCTGCTGCAAAATGAAAATTTCTGTCACAAAAAAGTATTAT 294
QY 181 TTTAAGAACTACCATCTCTGTTGGGTATGCCATTGGCAGACCTTTGACCTTACA 240
DB 295 TTTAAGAACTACCATCTCTGTTGGGTATGCCATTGGCAGACCTTTGACCTTACA 354
QY 241 TCCCTGCAAGCAATGTTCAATATCAAGGTGTCATCAACAAGACCGCTCATTTGAC 300
DB 355 TCCCTGCAAGCAATGTTCAATATCAAGGTGTCATCAACAAGACCGCTCATTTGAC 414
QY 301 AACAAATCCATGCGGTCTGATTCACCATAGAGACATAGCTGGATCTGACTAATCTTA 360
DB 415 AACAAATCCATGCGGTCTGATTCACCATAGAGACATAGCTGGATCTGACTAATCTTA 474
QY 361 CCTCAGCAGGCGGACCGCTTTCAGAAATGATTGGATGATTTAGAGTCACCGCACT 420
DB 475 CCTCAGCAGGCGGACCGCTTTCAGAAATGATTGGATGATTTAGAGTCACCGCACT 534
QY 421 CACACCCGCCAAAAGAGTGGCATTTGAACACTTTTCAACCTCACTTAATTCGCCGT 480
DB 535 CACACCCGCCAAAAGAGTGGCATTTGAACACTTTTCAACCTCACTTAATTCGCCGT 594
QY 481 GATTGAGATTCAGCAAGTGCCTTATAGCTCTTTCAGCGGTGAGCACAAAATCCCTTGTGTT 540
DB 595 GATTGAGATTCAGCAAGTGCCTTATAGCTCTTTCAGCGGTGAGCACAAAATCCCTTGTGTT 654
QY 541 GAAGTGCAGCAAGAGAGAGTGTGCTGGGTGTGAGTACTGAGACCTGAGACAT 600
DB 655 GAAGTGCAGCAAGAGAGAGTGTGCTGGGTGTGAGTACTGAGACCTGAGACAT 714
QY 601 GCCAGGTCACAGTATTTCAACGACCTCAGCAAGAAGTATGAAATCCACACCTATGGCAA 660

DB 715 GCCAGGTCACAGTATTTCAACGACCTCAGCAAGAAGTATGAAATCCACACCTATGGCAA 774
QY 661 GCATGCGAGATTAAGTGTAGCATTAATAAATGTTGCTCCACCATATCTACTGTAATTT 720
DB 775 GCATGCGAGATTAAGTGTAGCATTAATAAATGTTGCTCCACCATATCTACTGTAATTT 834
QY 721 TATCTTTTCATTGAAAACATTCATTCACAAAGATTACTACAGAAAAGCTCTACATGCA 780
DB 835 TATCTTTTCATTGAAAACATTCATTCACAAAGATTACTACAGAAAAGCTCTACATGCA 894
QY 781 TTTTGGCTGGTTCAGTACCTGTTGCTCCTGGGTCATTCAGGAAAACATAGAAATTAT 840
DB 895 TTTTGGCTGGTTCAGTACCTGTTGCTCCTGGGTCATTCAGGAAAACATAGAAATTAT 954
QY 841 ATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 900
DB 955 ATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1014
QY 901 TATCTGAGAGAGTTGACAAAAACAATTAAGTTTACTGTTACTGTTACTGTTACTGTTAG 960
DB 1015 TATCTGAGAGAGTTGACAAAAACAATTAAGTTTACTGTTACTGTTACTGTTACTGTTAG 1074
QY 961 GATTTTACGTAACTACCAAGCTTTTGGGAATCAGATGCTGCTGACATGACCAT 1020
DB 1075 GATTTTACGTAACTACCAAGCTTTTGGGAATCAGATGCTGCTGACATGACCAT 1134
QY 1021 GTAAAAAGCATCAAGAAATTAAGTCTGTTGTTGTTAGTATTTAGAGAAATGTTTGAATTA 1080
DB 1135 GTAAAAAGCATCAAGAAATTAAGTCTGTTGTTGTTAGTATTTAGAGAAATGTTTGAATTA 1194

RESULT 2
AF345993
LOCUS 1128 bp mRNA linear ROD 07-MAR-2001
DEFINITION Rattus norvegicus alpha1,3-fucosyltransferase IX (Fut9) mRNA,
complete cds.
ACCESSION AF345993
VERSION AF345993.1 GI:13242183
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1128)
Baboval,T., Henton,T., Klonally,E. and Smith,F.I.
Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX)
and comparison of the expression of fuc-TIV and fuc-TIX genes
during rat postnatal cerebellum development
J. Neurosci. Res. 62 (2), 206-215 (2000)
11020213
PUBMED 2 (bases 1 to 1128)
Smith,F.I. and Baboval,T.
REFERENCE
AUTHORS Smith,F.I. and Baboval,T.
TITLE Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver
JOURNAL Center, 200 Trapelo Rd., Waltham, MA, USA
FEATURES
source
1..1128
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1128
/gene="Fut9"
49..1128
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="AAK16591.1"
/db_xref="GI:13242184"
/translation="MTSTSKGILRPFLIVCIILGCFNACLLIYIKPTNSWVSPMESA
SSYLKMKNFPSKTDYFNETTILVWVFPQGFDTLSCAMFNIOGCHLTDLSLNK
SHAVLIHRDLSMDLNLPOQARPPPOKIMMNLSPHTPOKSGIELNLTLYR

DSDIQVPGFLVSTSPFVEVSKELVGVVSNMPEHAKVYIYNELSKSEIHTY
GQAFGEVNDKNIPISTCKEFLSENSHKOYITEKLYNALVASVPLPSREN
YENYIPADSFIVHEDNPSBELAKYLEKVDKNKLYLSFNMKDPVTNLPFWESHA
CIAODHVKRHOEKSVMNLEKFWFN"

BASE COUNT 324 a 256 c 228 g 320 t

ORIGIN

Query Match 93.3%; Score 1008; DB 10; Length 1128;
Best Local Similarity 95.8%; Pred. No. 5,7e-259;
Matches 1035; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATGACATCAACATCCAAAGCATTTCTGGCCATTTCTATCGCTGCATCATCTCGGC 60
DB 49 ATGACATCAACATCCAAAGCATTTCTGGCCATTTCTATCGCTGCATCATCTCGGC 108
QY 61 TGGTCATGGCATGTGCTCATTTATATCAAGCCACCAAGAGCTGGCTTCAGTCCA 120
DB 109 TGGTCATGGCATGTGCTCATTTATATCAAGCCACCAAGAGCTGGCTTCAGTCCA 168
QY 121 ATGAGATCTGCAAGTCTGTGCTGAAAAATGAAAAATTTCTTCACAAAACTGATTAT 180
DB 169 ATGAGATCTGCAAGTCTGTGCTGAAAAATGAAAAATTTCTTCACAAAACTGATTAT 228
QY 181 TTTAAGCAAACTACATTCCTGGTTGGGTATGGCCATTTGGCGAGACCTTTGACCTTACA 240
DB 229 TTTAAGTAAACTACTATTCTGGTTGGGTATGGCCATTTGGCGAGACCTTTGACCTTACA 288
QY 241 TCCGCGCAGCATGTTCATATATCAAGGGTGCATCTCAACAGACCGCTCATCTAC 300
DB 289 TCCGCGCAGCATGTTCATATCAAGGGTGCATCTCAACAGACCGCTCATCTAC 348
QY 301 AACAAATCCCATGCGGTCTGATTCACCATAGAGACATGCTGGATCTGACTACTTA 360
DB 349 AACAAATCCCATGCGGTCTGATTCACCATAGAGACATGCTGGATCTGACTACTTA 408
QY 361 CCTCAGAGAGCCAGCCACCTTTCCAGAAATGATTTGATGATTTAGAGTACCCACT 420
DB 409 CCTCAGAGAGCCAGCCACCTTTCCAGAAATGATTTGATGATTTAGAGTACCCACT 468
QY 421 CACACCCCCCAAAAAGATGAGCATTTGAACACTGTTCACCTGACTTAACTTACCGCGT 480
DB 469 CACACCCCCCAAAAAGATGAGCATTTGAACACTGTTCACCTGACTTAACTTACCGCGT 528
QY 481 GATTACAGATATCCAGATGCTTATGCTTCTTGAAGGTGAGCAAAATCCCTTTGTGTT 540
DB 529 GATTACAGATATCCAGATGCTTATGCTTCTTGAAGGTGAGCAAAATCCCTTTGTGTT 588
QY 541 GAAGTGCAGCAAGAGAGATGCTGCTGGTGTGAGTAACTGGAACCTGAGCAT 600
DB 589 GAAGTGCAGCAAGAGAGATGCTGCTGGTGTGAGTAACTGGAATCCCGAGCAT 648
QY 601 GCCAGGGTCAAGTATTACAGAGAGCTCAGCAAGAGTATGAATCCACACTTATGGCCA 660
DB 649 GCCAGGGTCAAGTATTACAGAGAGCTCAGCAAGAGTATGAATCCACACTTATGGCCA 708
QY 661 GCATTCGGAGATATCGTGAAGATGATTAATCTGATTCACCATATCTTACTTAAATTT 720
DB 709 GCATTCGGAGATATCGTGAAGATGATTAATCTGATTCACCATATCTTACTTAAATTT 768
QY 721 TATCTTTCATTTAAAACTAATTCACAAAGATTTACATCAGAAAAAGCTCTACATGCA 780
DB 769 TATCTTTCATTTAAAACTAATTCACAAAGATTTACATCAGAAAAAGCTCTACATGCA 828
QY 781 TTTTGGCTGGTTCAGTACCTGTTGCTGGTGCATCTGAGGAAAACTATGGAATTAAT 840
DB 829 TTTTGGCTGGTTCAGTACCTGTTGCTGGTGCATCTGAGGAAAACTATGGAATTAAT 888
QY 841 ATTCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
DB 889 ATTCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 948
QY 901 TATCTGAAGAAATGAGCAAAACATTAAGTGTACTTGTACTTAACTGAGAGAAAG 960

DB 949 TATCTGAAGAGGTGTGCAAAAAACATTAAGTTATATCTTACTTCAATCGAGAGAG 1008
QY 961 GATTTACTGTAACCTTACACAGCGTTTGGGAATACATGCATCGCTGAGTACCAT 1020
DB 1009 GATTTACTGTAACCTTACACAGCGTTTGGGAATACATGCATCGCTGAGTACCAT 1068
QY 1021 GTAAAAAGCATCAAGATATTAAGTCTGTTGGTATTAAGAAAGTTTGGAAATTA 1080
DB 1069 GTAAAAAGCATCAAGATATTAAGTCTGTTGGTATTAAGAAAGTTTGGAAATTA 1128

RESULT 3

AB049819

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

1..2156
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
346..1425
/gene="FUT9"
346..1425
/gene="FUT9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="BAB40953.1"
/db_xref="GI:13591589"
/translation="MSTSKGLRPFLVLCILGCVACLLIYIKPTNSWVSPMESA
SSVLRKKNFESKRTDFENETLLVWVPGGFDDLSCCAMNIGCHLITRSLYNK
SHAVLIHNRDIDMDLNLPOARPEPQKIMVNLSPPTPKSGIEHLFNLTYR
DSDIQVPGFLVSTSPFVEVSKELICVWVSNMPEHAKVYIYNELSKSEIHTY
GOAFGEVNDKNIPISTCKEFLSENSHKOYITEKLYNALVASVPLPSREN
YENYIPADSFIVHEDNPSBELAKYLEKVDKNKLYLSFNMKDPVTNLPFWESHA
CIAODHVKRHOEKSVMNLEKFWFN"

BASE COUNT 605 a 470 c 437 g 644 t

ORIGIN

Query Match 93.0%; Score 1004.8; DB 10; Length 2156;
Best Local Similarity 95.6%; Pred. No. 4.1e-258;
Matches 1033; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 ATGACATCAACATCCAAAGCATTTCTGGCCATTTCTATCGCTGCATCATCTCGGC 60
DB 346 ATGACATCAACATCCAAAGCATTTCTGGCCATTTCTATCGCTGCATCATCTCGGC 405
QY 61 TGGTCATGGCATGTGCTCATTTATATCAAGCCACCAAGAGCTGGCTTCAGTCCA 120
DB 406 TGGTCATGGCATGTGCTCATTTATATCAAGCCACCAAGAGCTGGCTTCAGTCCA 465

QY 121 ATGAGTGTGCAAGTCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 180
 DB 466 ATGAGTGTGCAAGTCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 525
 QY 181 TTTAAGAACTACCTATCTGCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 240
 DB 526 TTTAAGAACTACCTATCTGCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 585
 QY 241 TCTGCAAGCAAGTCTGCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 300
 DB 586 TCTGCAAGCAAGTCTGCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 645
 QY 301 AACAAATCCATGCGTCTGCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 360
 DB 646 AACAAATCCATGCGTCTGCTGCTGCGTGAATGAAAAATTTCTTCTCCACAAAACTGATAT 705
 QY 361 CCTCAGAGCCAGCCACCTTTGCAAAATGATGATGATGATGATGATGATGATGATGAT 420
 DB 706 CCTCAGAGCCAGCCACCTTTGCAAAATGATGATGATGATGATGATGATGATGATGAT 765
 QY 421 CACACCCCAAGAGTGCATTTGAACCTTTGCAAAATGATGATGATGATGATGATGATGAT 480
 DB 766 CACACCCCAAGAGTGCATTTGAACCTTTGCAAAATGATGATGATGATGATGATGATGAT 825
 QY 481 GATTCAGATATTCAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 540
 DB 826 GATTCAGATATTCAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 885
 QY 541 GAAGTGCAGAGCAAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 600
 DB 886 GAAGTGCAGAGCAAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 945
 QY 601 GCCAGGTCAGATTTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 946 GCCAGGTCAGATTTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
 QY 661 GCATTCGAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 720
 DB 1006 GCATTCGAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 1065
 QY 721 TATCTTCATTTGAAATTCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 780
 DB 1066 TATCTTCATTTGAAATTCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 1125
 QY 781 TTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 1126 TTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
 QY 841 ATTCAGCTGATTCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 900
 DB 1186 ATTCAGCTGATTCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 1245
 QY 901 TATCTGAAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 960
 DB 1246 TATCTGAAGAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 1305
 QY 961 GATTTTACGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 1306 GATTTTACGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1365
 QY 1021 GTTAAAGAGCAAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 1080
 DB 1366 GTTAAAGAGCAAGTGCATTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTTGAGCTTT 1425

RESULT 4
 AF230460 1705 bp mRNA linear ROD 12-JUL-2000
 LOCUS Cricetus griseus alpha(1.3)fucosyltransferase type IX (Fuc9)
 DEFINITION mRNA complete cds.
 ACCESSION AF230460
 VERSION AF230460.1 GI:9049663

KEYWORDS Chinese hamster.
 SOURCE Cricetus griseus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetus.
 REFERENCE 1 (bases 1 to 1705)
 AUTHORS Patnaik,S.K., Zhang,A., Shi,S. and Stanley,P.
 TITLE Alpha(1,3)fucosyltransferases expressed by the gain-of-function Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30 Arch. Biochem. Biophys. 375 (2), 332-332 (2000)
 JOURNAL MEDLINE 20166953
 PUBMED 10700388
 REFERENCE 2 (bases 1 to 1705)
 AUTHORS Patnaik,S.K., Shi,S. and Stanley,P.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of Medicine, Chanin 516, 1300 Morris Park Avenue, New York, NY 10461, USA
 FEATURES
 source location/Qualifiers
 1..1705 /organism="Cricetus griseus"
 /db_xref="taxon:10029"
 1..1705 /gene="Fuc9"
 /note="cgFuc9"
 62..1141 /gene="Fuc9"
 /note="Fuc-TIX"
 /codon_start=1
 /product="alpha(1,3)fucosyltransferase type IX"
 /protein_id="AA82412.1"
 /db_xref="GI:9049664"
 CDS
 1..1705 /translation="MTSTSGILRPLIVCVIIACFMACLIYIKPTNSWSPMSA
 SVLKKMFESTKTDYENETIIVWVPGQFEDLSCQAFNTIOGCHLTDRSLYK
 SHAULIHRDISMDLTNLPQOAPPOKIMNLSPPHPOKGIHLEFNLITVRR
 DSDIOVPYGLTVSTNPFVEVSKELICWVSNMNEHARVYVNELSKSEIHTY
 GOAFGEVYENKLIPTISTCKFYLSPENSIIHDYITELVNAFLAGSVPLVGPSEN
 YENYIPADSFTHEDYNSPELAKYKEVDKNNKLYLSYFMRKDFVNLPRWESHA
 CIACDHVRHDEYKSVGNLEKMFVN"
 BASE COUNT 515 a 336 c 320 g 534 t
 ORIGIN
 Query Match 90.2%; Score 974.4; DB 10; Length 1705;
 Best Local Similarity 93.9%; Pred. No. 5,6e-250;
 Matches 1014; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 1 ATGACATCAACATCCAAAGGCAATTCGCCCATTTCTTAATCCTTCATCATTCTGGGC 60
 DB 62 ATGACATCAACATCCAAAGGCAATTCGCCCATTTCTTAATCCTTCATCATTCTGGGC 121
 QY 61 TGCCTCATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 DB 122 TGCCTCATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
 QY 121 ATGAGAGTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 182 ATGAGAGTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 QY 181 TTTAAGCAAGTACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 242 TTTAAGCAAGTACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
 QY 241 TCTGCAAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 302 TCTGCAAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
 QY 301 AACAAATCCATGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 362 AACAAATCCATGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
 QY 361 CCTCAGAGCCAGCCACCTTTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 420

```

Db 422 CCTCAGCAGGCTAGGCCGCCCTTTGAGAAATGATTTGGATGAACCTAGAGTCTCCACT 481
QY 421 CACACCCCCCAAGAGTGGATGGAACACTTTGTTCAACTGACTCTACTTATCGCGT 480
Db 482 CATACCCCCCAAGAGTGGATGGAACACTTTGTTCAACTGACTCTACTTATCGCGT 541
QY 481 GATTGAGATATCCAAAGTGGCTTATGGCTTCTGACGGGTGAGCACAATTCCTTTGTGTT 540
Db 542 GATTGAGATATCCAAAGTGGCTTATGGCTTCTGACGGGTGAGCACAATTCCTTTGTGTT 601
QY 541 GAAGTGGCAGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
Db 602 GAAATGGCAGAGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 661
QY 601 GCCAGGGTCAAGTATTAACAAGCAGCAGAGATATGAAATCCACACTATGAGCAA 660
Db 662 GCCAGGGTCAAGTATTAACAAGCAGCAGAGATATGAAATCCACACTATGAGCAA 721
QY 661 GCATTCGAGAGATACGTAAGCAGATAAATCTGATTCGCCACATATCTACTTTAAATTT 720
Db 722 GCTTTGGAGATACGTAAGCAGATAAATCTGATTCGCCACATATCTACTTTAAATTT 781
QY 721 TATCTTCATTTGAAACCTCAATTCACAAAGATATACATCAGAGAAAGCTCTAATGCA 780
Db 782 TATCTTCATTTGAAACCTCAATTCACAAAGATATACATCAGAGAAAGCTCTAATGCA 841
QY 781 TTTTGGCTGCTCAGTACGACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 842 TTTTGGCTGCTCAGTACGACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
QY 841 ATTCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 900
Db 902 ATTCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 961
QY 901 TATCTGAGAGATGTCACAAACAAATATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 962 TATCTGAGAGATGTCACAAACAAATATGATGATGATGATGATGATGATGATGATGATGAT 1021
QY 961 GATTTCATCTAAACCTACACAGGTTTGGGAATCAGATGATGATGATGATGATGATGATGAT 1020
Db 1022 GATTTCATCTAAACCTACACAGGTTTGGGAATCAGATGATGATGATGATGATGATGATGAT 1081
QY 1021 GTAAAAAGGATCAAGATATTAAGTCTGTTGGTAATTTAGAGAAATGTTTGGAAATTA 1080
Db 1082 GTAAAAAGGATCAAGATATTAAGTCTGTTGGTAATTTAGAGAAATGTTTGGAAATTA 1141

RESULT 5
LOCUS HSA238701 2501 bp mRNA linear PRI 21-OCT-2000
DEFINITION Homo sapiens mRNA for alpha-3-fucosyltransferase.
ACCESSION AJ238701
VERSION AJ238701.1 GI:4741566
KEYWORDS alpha-3-fucosyltransferase; FUT9 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2501)
AUTHORS Callican-Thomas,A., Coullin,P., Candellier,J.J., Balanzino,L.,
Mennesson,B., Oriol,R. and Mollicone,R.
TITLE FUT9 and FUT9 genes are expressed early in human embryogenesis
JOURNAL Glycobiology 10 (8), 789-802 (2000)
MEDLINE 20386991
REFERENCE 2 (bases 1 to 2501)
AUTHORS Mollicone,R.
TITLE Direct Submission
COMMENT Submitted (29-APR-1999) Mollicone R., Glycobiology, INSERM U504, 16
AV. Paul Vaillant-Couturier, Villejuif, 94807 Cedex, FRANCE
FEATURES
location/Qualifiers
1..2501
source /organism="Homo sapiens"

```

```

/gb_xref="taxon:9606"
/chromosome="6"
/tissue_type="whole embryo"
/dev_stage="40 to 60 days"
/gene="FUT9"
/gene="FUT9"
105..1184
/gene="FUT9"
/codon_start=1
/product="alpha-3-fucosyltransferase"
/protein_id="CA41890.1"
/db_xref="GI:4741567"
/db_xref="SPTREMBL:09231"
/translation="MTSTSKGILRPLFVLCIILGCFMACLLIYIKPTNSWIFSPMSA
SSYLKMKNEPSTKDYENETITLWVWPEQGFDTLSCAMFNIQGHITLDRSLYK
SHAVLIHHDLSMDLNLROARPPROKTMNLESPTTPQSGIEHFNULTYR
DSDIQVPLGFLTYSTNPEVPSKELVCMVSNMPEHARKYINLSKSIIEHTY
GDAFGEVNDKNLIPITISACKFYLSEKILHKOYITEKLYNAELASVPLVPSREN
YENYIPADSEIHVEDYNSPELAKLYLEVDKNKLYLSYFNMRKDETVNLPRWESHA
CLACDHYKRRQEKSVNLEKFWN"
2192
/gene="FUT9"
/evidence=experimental
2235
/gene="FUT9"
/evidence=experimental
2310
/gene="FUT9"
/evidence=experimental
2466
/gene="FUT9"
/evidence=experimental
BASE COUNT 818 a 436 c 422 g 825 t
ORIGIN
Query Match 86.7%; Score 936; DB 9; Length 2501;
Best Local Similarity 91.7%; Pred. No. 1.1e-239;
Matches 990; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 1 ATGACATCAACATCCAAAGCATCTTCGCCCATTTCTAATCGTCGATCATCTCGGCG 60
Db 105 ATGACATCAACATCCAAAGCATCTTCGCCCATTTCTAATCGTCGATCATCTCGGCG 164
QY 61 TGCCTCATGCGATGTCGTCATTTATATCAAGCCCAACACAGCTGGGCTTCAGTCA 120
Db 165 TGTTCATGCGATGTCGTCATTTATATCAAGCCCAACACAGCTGGGCTTCAGTCA 224
QY 121 ATGAGTCTGCAAGTCTGCTGGAATGAAATGAAATTTCTTCACAAAAATGATTA 180
Db 225 ATGGAATCAGCAGACTCTGCTGAAATGAAATGAAATTTCTTCACAAAAATGATTA 284
QY 181 TTTAACAAGATCAATTCGATTTGGGTATGAGCATTTGGCGAGACTTTGACCTTACA 240
Db 285 TTTAATATAACTACTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
QY 241 TCTGCGCAGCAATGTTCAATATCAAGGGTGCATCTCAACAAAGACGCGTCACTGAT 300
Db 345 TCTGCGCAGCAATGTTCAATATCAAGGGTGCATCTCAACAAAGACGCGTCACTGAT 404
QY 301 AACCAATCCATGCGGCTGCTGATTCACCAATAGATGATTTAGATGATGATGATGAT 360
Db 405 AACCAATCCATGCGGCTGCTGATTCACCAATAGATGATTTAGATGATGATGATGAT 464
QY 361 CCTAGCAGAGCGAGCCACCTTTCAAAATAGATTTGGATGATTTAGATGATGATGAT 420
Db 465 CCTAGCAGAGCGAGCCACCTTTCAAAATAGATTTGGATGATTTAGATGATGATGAT 524
QY 421 CACACCCCCCAAGAGTGGATGGAACACTTTGTTCAACTGACTCTACTTATCGCGT 480
Db 525 CACACCCCCCAAGAGTGGATGGAACACTTTGTTCAACTGACTCTACTTATCGCGT 584
QY 481 GATTGAGATATCCAAAGTGGCTTATGGCTTCTGACGGGTGAGCACAATTCCTTTGTGTT 540

```



```
repeat_region 13168..13280
/note="LIMAB repeat: matches 5710..5832 of consensus"
repeat_region 13851..17218
/note="LIPB1 repeat: matches 2787..6155 of consensus"
repeat_region 18903..18953
/note="MIR2CB repeat: matches 439..501 of consensus"
repeat_region 18984..19433
/note="MIR2CB repeat: matches 1..460 of consensus"
/note="match: (19108..19690)
/note="match: GSS: Em:AQ563935"
20683..20801
/note="AluIo/FRAM repeat: matches 169..288 of consensus"
20802..20862
/note="AluIo/FLAM repeat: matches 3..66 of consensus"
20912..21033
/note="MIR repeat: matches 114..237 of consensus"
21047..21128
/note="41 copies 2 mer ta 63% conserved"
21676..21855
/note="L2 repeat: matches 2130..2308 of consensus"
21890..23612
/note="LIP repeat: matches 841..2570 of consensus"
23601..26609
/note="LIPAS repeat: matches 3107..6145 of consensus"
26620..26789
/note="MER5A repeat: matches 1..174 of consensus"
28521..28627
/note="MER39 repeat: matches 5..111 of consensus"
28656..28837
/note="MER39b repeat: matches 54..235 of consensus"
28838..29142
/note="AluSc repeat: matches 4..308 of consensus"
29143..29494
/note="MER39b repeat: matches 235..549 of consensus"
30529..31344
/note="LIPAS repeat: matches 5331..6146 of consensus"
32234..32340
/note="L2 repeat: matches 2627..2750 of consensus"
32399..32578
/note="MIR repeat: matches 12..203 of consensus"
32816..32979
/note="MIR repeat: matches 91..259 of consensus"
33312..33596
/note="LIM4 repeat: matches 3473..3773 of consensus"
34856..34915
/note="30 copies 2 mer aa 68% conserved"
/note="match: (34987..35547)
/note="match: GSS: Em:AQ321598
match: STS: Em:G55571"
complement(36488..36936)
/note="match: GSS: Em:B41073"
complement(36654..37090)
/note="match: GSS: Em:AQ772584"
37102..37591
/note="match: GSS: Em:AQ489567"
complement(37581..38221)
/note="match: GSS: Em:AQ418233"
37720..37772
/note="L2 repeat: matches 2657..2710 of consensus"
complement(37778..38215)
/note="match: GSS: Em:AQ819773"
38218..38732
/note="match: GSS: Em:AQ599868"
38861..39164
/note="AluSc repeat: matches 1..305 of consensus"
39612..39653
/note="21 copies 2 mer ac 90% conserved"
40854..40905
/note="L2 repeat: matches 2655..2710 of consensus"
```

```
repeat_region 42169..42270
/note="MER5A repeat: matches 1..111 of consensus"
repeat_region 42411..42548
/note="MER5B repeat: matches 21..181 of consensus"
repeat_region 42780..42938
/note="MER5B repeat: matches 17..177 of consensus"
repeat_region 43229..43326
/note="MIR repeat: matches 40..148 of consensus"
45383..45428
/note="MER5B repeat: matches 18..62 of consensus"
complement(47284..47971)
/note="match: STS: Em:G51955
match: GSS: Em:AQ269059"
complement(47537..48011)
/note="match: STS: Em:G36899"
complement(47723..48014)
/note="match: STS: Em:G30049"
47972..48451
/note="match: GSS: Em:AQ36527"
complement(48518..48985)
/note="match: GSS: Em:AQ801834"
48597..48803
repeat_region

Query Match 86.4% Score 932.8; DB 9; Length 80247;
Best Local Similarity 91.5% Pred. No. 8.5e-239;
Matches 988; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGCATCAACATCAAGGATCTTGGCCATTCGTAATCGTCGATCAATCCGGC 60
|||||
DB 35559 ATGCATCAACATCAAGGATCTTGGCCATTCGTAATCGTCGATCAATCCGGC 35618

QY 61 TGTCTATGGCATGTCTGCTCATTTATATCAAGCCCAACACAGCTGGCTTCAGTCA 120
|||||
DB 35619 TGTCTATGGCATGTCTGCTCATTTATATCAAGCCCAACACAGCTGGCTTCAGTCA 35678

QY 121 ATGGAGTCTGCAAGTCTGTGCTGAAATGAAATTTCTTCCACAAAATGATAT 180
|||||
DB 35679 ATGGAGTCTGCAAGTCTGTGCTGAAATGAAATTTCTTCCACAAAATGATAT 35738

QY 181 TTTATAGCAACTACATTCGTTGGTGTATGCGATTTGGGCGACCTTGACCTTACA 240
|||||
DB 35739 TTTATAGCAACTACATTCGTTGGTGTATGCGATTTGGGCGACCTTGACCTTACA 35798

QY 241 TCTCGCAAGCAATGTTCAATATCAAGGGTGCATCTCACACAGCCGTCATGTATC 300
|||||
DB 35799 TCTCGCAAGCAATGTTCAATATCAAGGGTGCATCTCACACAGCCGTCATGTATC 35858

QY 301 AACCAATCCATGGCTGCTGATTCACCATAGACATCAGCTGGATCTGACTACTTA 360
|||||
DB 35859 AACCAATCCATGGCTGCTGATTCACCATAGACATCAGCTGGATCTGACTACTTA 35918

QY 361 CCTGAGCAGGCGACGCCCTTCCAGAAATGATTTGGATTTAGAGTCCACCACT 420
|||||
DB 35919 CCTGAGCAGGCGACGCCCTTCCAGAAATGATTTGGATTTAGAGTCCACCACT 35978

QY 421 CACACCCCAAGAGATGGATTAACACATTTGTTCAACCTGACTTAATTCGCGCT 480
|||||
DB 35979 CACACCCCAAGAGATGGATTAACACATTTGTTCAACCTGACTTAATTCGCGCT 36038

QY 481 GATTGAGATATCCAGTGCCTTATGCTTCTTGACGCTGACACAAATCCCTTTGTTT 540
|||||
DB 36039 GATTGAGATATCCAGTGCCTTATGCTTCTTGACGCTGACACAAATCCCTTTGTTT 36098

QY 541 GAAGTGCAGCAAGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
|||||
DB 36099 GAAGTGCAGCAAGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 36158

QY 601 GCCAGGCTCAAGTATTAACAAGCTCAGCAAGATTTAAATCCACACTATGGCAA 660
|||||
DB 36159 GCCAGGCTCAAGTATTAACAAGCTCAGCAAGATTTAAATCCACTACTAGGGCAA 36218

QY 661 GCATTGGAGATATCGTGAACGATAAATCTGATTCACACCATATCTACTTAAATTT 720
|||||
```


BASE COUNT	937 a	594 c	541 g	947 t	ORIGIN
	CLACDHYKXHOEYKXSVGLSEKMEWN"				
Query Match	86.2%; Score 931.2; DB: 9; Length 3019;				
Best Local Similarity	91.4%; P: 2.1e-288;				
Matches 907; Conservative	0; Mismatches 93; Indels 0; Gaps				
QY	1	ATGACATCAACATCCAAAGCATTTCTGGCCATTTCTATATGCTGTGCATATCTGGGC	60		
DB	295	ATATACATCAACATCCAAAGGATTTCTGGCCATTTTAAATGTGTGCATATTTCTGGGC	354		
QY	61	TGCTTCATGCAATGTCGTCTCATTTTATCAAGCCCAACCAAGCTGGGTCTTGAGTCA	120		
DB	355	TGTTTATAGCCATGCTCTTCATATTTACATCAAAACCTACCAACAGCTGGATCTGAGTCA	414		
QY	121	ATGAGCTGTGCAAGTTCGTGTCTGTAATAAAGAAAAATTTCTCTCCACAAAACTATATAT	180		
DB	415	ATGAGATTCAGCCAGCTGTGTCTGTAATAAAGAAAAATTTCTCTCCACAAAACTATATAT	474		
QY	181	TTTTAAGCAACATCACTTCGTGTGGGTATGGCCATTTGGGGAGACCTTTACCTTACCA	240		
DB	475	TTTTAAGCAACATCACTTCGTGTGGGTATGGCCATTTGGGGAGACCTTTGACCTTACA	534		
QY	241	TCTGTCAAGCAATGTTCATATATCCAAAGGTGCATTCACACACAGCCGTCACTTATAC	300		
DB	535	TCTGTCAAGCAATGTTCACATATCCAAAGGTGCATTCACACAGCCGTTCACGTATAC	594		
QY	301	AACAAATCCCATGGGTCGTGTGATTCACACATAGAGACATCACTGGGTATGCACTTAA	360		
DB	595	AACAAATCCCATGCACTTGTGTGATTCACACAGACATCACTTGGGTATGCACTTAA	654		
QY	361	CGTCACAGCCAGCCAGCCCTTTCAGAAATGGATTTGATTTAGATGACCCACT	420		
DB	655	CGTCACAGCCAGCCAGCCCTTTCAGAAATGGATTTGATTTAGATTTAGATGACCCACT	714		
QY	421	CACACCCCCCAAAAGATGGCATTTGAACACTTGTTCACACTGACTTAACTTATGGCGT	480		
DB	715	CACACCTCCCCCAAAAGAGTGATTTGACACTTGTTTAACTGTGACTGTGATTCACCGCT	774		
QY	481	GATTCAGATATCCAAATGTCCTTATGCTTCTTACGCTGAGACACAAATCCCTTGTGTTT	540		
DB	775	GATTCAGATATCCAAATGTCCTTATGCTTCTTGAAGGTACACAAATCCCTTGTGTTT	834		
QY	541	GAAGTCCCAAGCAGAGAAAGTTGTGCTGTGGTTGTGATTAATCTGAAACCTGACAT	600		
DB	835	GAAGTCCCAAGCAGAGAAATTTGTGCTGTGGTTGTGATTAATCTGAAACCTGACAT	894		
QY	601	GGCAGGCTCAAGTATTTACACAGAGCTCACACAGATATTGAAATCCACACCTGTGGCAA	660		
DB	895	GGCAGAGCTCAAGTATTTACAAATGAGCTTACGAAACAAATGAAATCCATACCTGTGGCAA	954		
QY	661	GCATTCGGGAAATACGTACGATATAAAAAATTCGATTTCCACCATATATCTTAAATTTT	720		
DB	955	GCATTTGGGAAATATGATATGATATAAAAAATTTGATTTCTACCATATGTGTTGAAATTT	1014		
QY	721	TATCTTCATTTTGAAGAACTCAATTCACAAAGATTTTACATACAGAGAAAGCTTACCAATCA	780		
DB	1015	TATCTTCATTTTGAAGAAATTCATATCCACAGATTTACATACAGAGAAAGCTTATCAATGCT	1074		
QY	781	TTTTTGGCTGTGTACGTAATCTGTTCCTGTGGCTCATATAGGAGAAACTTATGAGATTTAT	840		
DB	1075	TTTTTGGCTGTGTGTACGTAATCTGTTCCTGTGGGACCATATAGGAGAAACTTATGAGATTTAT	1134		
QY	841	ATTTCAGCTGATTCATTCATGTGAGAAATTTTAACTCTCCACAGTAGATTAGCAAA	900		
DB	1135	ATTTCAGAGATTCATTCATGTGAGAAATTTTAACTCTCCACAGTAGAGTAGCAAA	1194		
QY	901	TATCTGAGAGAGTTGACAAAAACATTAAGTTGTACCTTATGTTACTTTAACTGGAGAAAG	960		
DB	1195	TATCTGAGAGAGTTGACAAAAACATTAAGTTATACCTTATGTTACTTTAACTGGAGAGAG	1254		
QY	961	GATTTTACGTAAACCTACAGGGTTTTGGGATACATATGATCCTGTGATCGAGCAT	1020		

OY	1021	GTAAGGACATCAAGAATAATAGTCTGTGGTAAATTAGAGAAAATGGTTTGGAATTAA	1080
Db	1315	GTGAAGAGCATCAGAAATATATAGTCGTGTGTAAATTAGAGAAATGGTTTGGAATTAA	1374
RESULT 8			
LOCUS	AC100674		
DEFINITION	Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.		
ACCESSION	AC100674		
VERSION	AC100674.1 GI:17048040		
KEYWORDS	HTG; HTGS; PHASEO.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.		
AUTHORS	Birken,B., Linton,L., Nusbaum,C. and Lander,E.		
TITLE	1 (bases 1 to 71396)		
JOURNAL	Mus musculus, clone RP23-167K24		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 71396)		
	Birken,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarita,J., Campdiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collumore,A., Cook,A., Cooke,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Glade,S., Goette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J., Menusz,L., Mihova,T., Mlenka,V., Murphy,T., Naylor,U., Nguyen,C., Nobdu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhkhap,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnuppak,R., Seaman,S., Severly,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J.J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zemke,l., Zimmer,A. and Zody,N.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
COMMENT	All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html		
	----- Genome Center ----- Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence.submissions@genome.wi.mit.edu ----- Project Information ----- Center project name: L15932 Center clone name: 167_K_24 ----- * NOTE: This record contains 71 individual * sequencing reads that have not been assembled into * contigs. Runs of N are used to separate the reads * and the order in which they appear is completely * arbitrary. Low-pass sequence sampling is useful for * identifying clones that may be gene-rich and allows * overlap relationships among clones to be deduced. * However, it should not be assumed that this clone * will be sequenced to completion. In the event that * the record is updated, the accession number will * be preserved. 1 788: contig of 788 bp in length		

789	888:	gap of	100 bp	
889	1702:	contlg of 814 bp	in length	
1703	1802:	gap of	100 bp	
1803	2642:	contlg of 840 bp	in length	
2643	2744:	gap of	100 bp	
2743	3631:	contlg of 895 bp	in length	
3638	3737:	gap of	100 bp	
3738	4633:	contlg of 898 bp	in length	
4636	4735:	gap of	100 bp	
4736	5651:	contlg of 922 bp	in length	
5658	5757:	gap of	100 bp	
5758	6692:	contlg of 935 bp	in length	
6693	6792:	gap of	100 bp	
6793	7703:	contlg of 911 bp	in length	
7704	7803:	gap of	100 bp	
7804	8742:	contlg of 939 bp	in length	
8743	8842:	gap of	100 bp	
8843	9787:	contlg of 945 bp	in length	
9788	9887:	gap of	100 bp	
9888	10818:	contlg of 932 bp	in length	
10820	10919:	gap of	100 bp	
10920	11808:	contlg of 889 bp	in length	
11809	11908:	gap of	100 bp	
11909	12830:	contlg of 922 bp	in length	
12831	12930:	gap of	100 bp	
12931	13847:	contlg of 917 bp	in length	
13848	13947:	gap of	100 bp	
13948	14866:	contlg of 939 bp	in length	
14867	14966:	gap of	100 bp	
14967	15911:	contlg of 925 bp	in length	
15912	16011:	gap of	100 bp	
16012	16903:	contlg of 892 bp	in length	
16904	17003:	gap of	100 bp	
17004	17906:	contlg of 903 bp	in length	
17907	18006:	gap of	100 bp	
18007	18931:	contlg of 925 bp	in length	
18932	19031:	gap of	100 bp	
19032	19945:	contlg of 914 bp	in length	
19946	20045:	gap of	100 bp	
20046	20951:	contlg of 906 bp	in length	
20952	21051:	gap of	100 bp	
21052	21949:	contlg of 898 bp	in length	
21950	22049:	gap of	100 bp	
22050	22950:	contlg of 901 bp	in length	
22951	23050:	gap of	100 bp	
23051	23945:	contlg of 895 bp	in length	
23946	24045:	gap of	100 bp	
24046	24955:	contlg of 910 bp	in length	
24956	25055:	gap of	100 bp	
25056	25960:	contlg of 925 bp	in length	
25961	26080:	gap of	100 bp	
26081	27052:	contlg of 972 bp	in length	
27053	27152:	gap of	100 bp	
27153	28067:	contlg of 915 bp	in length	
28068	28167:	gap of	100 bp	
28168	29076:	contlg of 909 bp	in length	
29077	29166:	gap of	100 bp	
29177	30093:	contlg of 917 bp	in length	
30094	30193:	gap of	100 bp	
30194	31099:	contlg of 906 bp	in length	
31100	31199:	gap of	100 bp	
31200	32097:	contlg of 898 bp	in length	
32098	32197:	gap of	100 bp	
32198	33113:	contlg of 916 bp	in length	
33114	33213:	gap of	100 bp	
33214	34125:	contlg of 912 bp	in length	
34126	34225:	gap of	100 bp	
34226	35133:	contlg of 908 bp	in length	
35134	35223:	gap of	100 bp	
35234	36150:	contlg of 917 bp	in length	
36151	36280:	gap of	100 bp	
36281	37171:	contlg of 921 bp	in length	
37172	37271:	gap of	100 bp	

37272	38159:	contig of 888 bp	in length
38160	38259:	gap of 100 bp	
38260	38156:	contig of 897 bp	in length
39257	39256:	gap of 100 bp	
39257	40188:	contig of 937 bp	in length
40189	40288:	gap of 100 bp	
40289	41210:	contig of 922 bp	in length
41211	41310:	gap of 100 bp	
41311	44222:	contig of 912 bp	in length
42323	42322:	gap of 100 bp	
43233	43218:	contig of 896 bp	in length
43319	43318:	gap of 100 bp	
43319	44233:	contig of 915 bp	in length
44234	44333:	gap of 100 bp	
44334	45246:	contig of 913 bp	in length
45247	45346:	gap of 100 bp	
45347	46261:	contig of 915 bp	in length
46262	46361:	gap of 100 bp	
46362	47277:	contig of 918 bp	in length
47280	47379:	gap of 100 bp	
47380	48276:	contig of 897 bp	in length
48277	48377:	gap of 100 bp	
48377	49283:	contig of 906 bp	in length
49283	49389:	gap of 100 bp	
49389	50305:	contig of 923 bp	in length
50306	50405:	gap of 100 bp	
50406	51314:	contig of 909 bp	in length
51315	51414:	gap of 100 bp	
51415	52333:	contig of 921 bp	in length
52336	52433:	gap of 100 bp	
52436	53163:	contig of 728 bp	in length
53164	53263:	gap of 100 bp	
53264	54210:	contig of 947 bp	in length
54211	54310:	gap of 100 bp	
54311	55199:	contig of 889 bp	in length
55200	55299:	gap of 100 bp	
55300	56212:	contig of 913 bp	in length
56213	56313:	gap of 100 bp	
56313	57287:	contig of 975 bp	in length
57288	57387:	gap of 100 bp	
57388	58245:	contig of 856 bp	in length
58246	58346:	gap of 100 bp	
58346	59235:	contig of 891 bp	in length
59237	59336:	gap of 100 bp	
59337	60272:	contig of 936 bp	in length
60273	60372:	gap of 100 bp	
60373	61289:	contig of 917 bp	in length
61290	61389:	gap of 100 bp	
61390	62299:	contig of 910 bp	in length
62300	62399:	gap of 100 bp	
62400	63322:	contig of 923 bp	in length
63323	63422:	gap of 100 bp	
63423	64336:	contig of 914 bp	in length
64337	64436:	gap of 100 bp	
64437	65355:	contig of 918 bp	in length
65355	65454:	gap of 100 bp	
65455	66355:	contig of 898 bp	in length
66355	66455:	gap of 100 bp	
66455	67365:	contig of 910 bp	in length
67366	67465:	gap of 100 bp	
67466	68370:	contig of 908 bp	in length
68371	68470:	gap of 100 bp	
68471	69379:	contig of 909 bp	in length
69380	69479:	gap of 100 bp	

Query Match	68.5%;	Score 740;	DB 2;	Length 71396;
Best Local Similarity	83.48;	Pred. No. 3.4e-187;		
Matches 844;	Conservative	0;	Mismatches 154;	Indels 14;
				Gaps 6

0y 72 AAGTCGTGCATTATATACAAAGCCACCACAAGCTGGGCTTCAGTCCAATGGAGTCMGC 131
||| | | | ||||||||| ||||||||||||
Db 65456 ATCTCTGCCCCCTGTGGTGGAATTCACCAACAAGACTCGGCTTTCAAGTCCAAATGGAGTCMGC 65515

[illegible]

LOCUS	AB035906	1707 bp	DNA	linear	VRT 05-JAN-2002
DEFINITION	Gallus gallus gene for cFtn9, complete cds.				
ACCESSION	AB035906				
VERSION	AB035906.1	GI:18146665			
KEYWORDS	cFtn9.				
SOURCE	Gallus gallus brain DNA.				

ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (sites)
AUTHORS Kaneo,M., Nishihara,S., Kitano,T., Nariyatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1707)
AUTHORS Kaneo,M., Saitou,N. and Kitano,T.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1999) Mika Kaneo, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Yata 1111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6789)
COMMENT Sequence updated (29-Feb-2000).
FEATURES
source
1..1707
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="Brain"
268..1347
/gene="cfuc-TIX"
268..1347
/gene="cfuc-TIX"
/codon_start=1
/product="CFU9"
/protein_id="BAB82489.1"
/db_xref="GI:18146866"
/translation="MTSTSKGIFRPFFIVLGCFMALIIYIKPTSSWISNPESA
SSVLKMKSEFFSKDTNENETILVLMPEQGFEDLSCQTMNIGHCHLIDRLSYNK
SHAVLIHHRDINDMDLNLPOARPEOKIMNLSPTPKRSKIEHLSFNLTYR
DSIDIOVYGMIVGATFYEVSKENLVCVMVSNMNPENKRYNLSKYLEHTY
GOAFGQVYNDKMLIPITISCKEFLSPENSIHDDYTERKYNLLAGSVPVYGPSEN
YENYIPADSFIVEDEPLSPRELAAYLIMLDKNNKATLSFNNKRPVSHLPFWESHA
CLACDHYRHQKTSIGINLEKFWN"
BASE COUNT 517 a 338 c 329 g 521 t 2 others
ORIGIN
Query Match 67.4%; Score 727.6; DB 5; Length 1707;
Best local Similarity 79.5%; Pred. No. 6.5e-184;
Matches 859; Conservative 1; Mismatches 220; Indels 0; Gaps 0;
QY 1 ATGACATCAACATCAAGGATCTTCGCCCATTTCTAATCGTCGATCATCTGGGC 60
DB 268 ATGACATCAACATCAAGGATCTTCGCCCATTTCTAATCGTCGATCATCTGGGC 327
QY 61 TCGTTCATGCGATGTCTGCTCATTTATATCAAGCCCAACAGCGTGGCTTCAGTCCA 120
DB 328 TGTTCATGCGATTTACTAATTTATTAATTAACCAACAGCGTGGATCTCTAATCCT 387
QY 121 ATGAGCTGCAAGTCTGCGCGAAGAAATTTCTTCCTCCCAAAACTGATTAAT 180
DB 388 GTGGAATCAAGCGTCAAGTTTGAAGAAATGAAAGCTCTTCTTCCAAAACTGATTAAT 447
QY 181 TTTAAGAACTACATCTTCGTTGGTATGGCCATTTGGGAGACCTTTGACTTACA 240
DB 448 TTAAATGAAGATTAATTTGGTTGGCTTGGCCATTTGGCCAGACATTCGATCTCAAC 567
QY 241 TCTGCGAAGAGATTTCAATATCAAGGATGCGATCTCAACAGACGCTCATTTGAC 300
DB 508 TCTGCGAAGAGATTTCAACATCAAGGATGCGATCTCAACAGCTCATCTCACTATAT 567
QY 301 AACAAATCCATGCGGTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 360
DB 568 AACAAATCCATGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 627
QY 361 CCTCAGACAGGCGACGCCCTTTCAGAAATGATTTGATGATTTAGATTCAGACCCACT 420
DB 628 CCTCAGACAGGCGACGCCCTTTCAGAAATGATTTGATGATTTAGATTCAGACCCACT 687
QY 421 CACACCCCAAAAAGATGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 480
DB 688 CATATCTCCACAAAAGATGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 747
QY 481 GATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 747
DB 748 GATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 748
QY 541 GAAATGCCAAGCAAGGAGAAATTTGATGCTGGGTTGTGATGATTCAGATTCAGATTCAGATTCAGATTC 749
DB 808 GAAATGCCAAGCAAGGAGAAATTTGATGCTGGGTTGTGATGATTCAGATTCAGATTCAGATTCAGATTC 750
QY 601 GCCAGGCTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 751
DB 868 GCTTCGAGTCAAGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 752
QY 661 GCATTCGAGTCAAGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 753
DB 928 GCTTCGAGTCAAGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 754
QY 721 TATCTTCATTTGAAAGTCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 755
DB 988 TATCTTCATTTGAAAGTCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 756
QY 781 TTTTGGCTGCTGCTCAGTACCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
DB 1048 CTCTGCT 758
QY 841 ATTCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 759
DB 1108 ATTCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 760
QY 901 TATCTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 761
DB 1168 TATCTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 762
QY 961 GATTTTACTGTTAACTACACGCTTTGGGAAATCAATGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 763
DB 1228 GATTTTACTGTTAACTACACGCTTTGGGAAATCAATGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 764
QY 1021 GTAAAGAGCATCAAGATTAATTAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1080
DB 1288 GTAAAGAGCATCAAGATTAATTAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1081
RESULT 10
AB035905 1080 bp DNA linear VRT 05-JAN-2002
LOCUS AB035905
DEFINITION Xenopus laevis gene for xFUT9, complete cds.
ACCESSION AB035905
VERSION AB035905.1 GI:18146863
KEYWORDS xFUT9.
SOURCE
ORGANISM
Xenopus laevis DNA.
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
1 (sites)
AUTHORS Kaneo,M., Nishihara,S., Kitano,T., Nariyatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1080)
AUTHORS Kaneo,M., Saitou,N. and Kitano,T.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1999) Mika Kaneo, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Yata 1111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6789)
COMMENT Sequence updated (29-Feb-2000).
FEATURES
source
1..1080
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/gene="xfuc-TIX"


```

Db      461 TCAGAAACCAATCTTGTATATGCGGTATCTTTTCGAGACAGCTTTTACCTGGAGCT 520
Oy      243 CTGCCAAGCAATGTTCATATATTCAGAGGTCCTTCACAGACAGCCGCTCATTTGACAA 302
Db      521 GTGTGGTGGCACTTCAATCCAGCGCTGTCTATTTAAGACAGACAGAAATGATGATCA 580
Oy      303 CAATTCACATGCGGCTTCATTCACCATAGACATCACTGGGATCTGATTAACCTTACC 362
Db      581 GAAAGCGCATGGAGTTATGTTTACACAGAGATCTGACGGGAT -----TTGCC 631
Oy      363 TCAGACAGCGCCAGCCACCTTTCAGAAATGATTTGGATGAAATTTAGATCCACCTCA 422
Db      632 GCAACCCCTTCAGCAGCTTTCACAAAGGATATGGTGAATATGAGATCTCCGATGA 691
Oy      423 CACCCCCCAAAAGATGGCATTAACCTTTCACTGACTCTAACCTTATGCGCGTGA 482
Db      692 TTCCTCCAAACAGCTTGTATTAATGATGGGTTTAACTGACGTCAAGTTTCGCAAGA 751
Oy      483 TTCAGATATCAAGTGCCTTA--TGGCTTCTGACGCTGACACAAATCCCTTGTGTT 539
Db      752 TTCAGATATTCAGAGCCTTACGGTGACATCATGATGCCACGAGAAAGAAACTT 811
Oy      540 TGAAGTGCACAGCAGAGAAAGTGTGTGCTGGTGGTGGATTAAGTCAACCTGAGCA 599
Db      812 CACCAATCCCAAAAGAGATTAACCTGTTGCTGATGATGATTAACCTTCAAGAACACCA 871
Oy      600 TGCAGGGTCAAGTATTACACAGCTCAGCAAGATTTGAATTCACACCTATGGCA 659
Db      872 CAAAGCTTCAAGTACTATTAATGAGTGTGTGAACACATTTAAAGTGAGAGCTATGAG 931
Oy      660 AGCAATTCGAGAAATACGTGACAGATAAAAATCTGATCCACCATATCTACTTGTAAAT 719
Db      932 GCACCTTCAATACCGGGTTAATGATGACGATTATGGAATGTGATCCAGTTGTAAT 991
Oy      720 TTATCTTCAATTTGAAACTCAATTCACAAAGATTTACATCAGAGAAAGCTTACATGC 779
Db      992 CTACCTTCTCTTGGAGAACTCCATTCACAGACATATTTCACTGAGAAAGCTTCAATCC 1051
Oy      780 ATTTTGGCTGTGAGTACTGCTGTGCTGCTGCTGCTATAGGAAATATAGAAATTA 839
Db      1052 TTTAGCGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1111
Oy      840 TATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 899
Db      1112 TATACCAAGACGCGCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1171
Oy      900 ATATCTGAAGCAAGTTGACAAAAACATATAGTTTACTTACTTACTTACTTACTTACTT 959
Db      1172 TCACCTTAATCTTGGACCAAAAGAGACAGATATTCCTCACTGAGAA 1231
Oy      960 GGATTTTACTGTAACCTTCACGCGTTTGGAGATCAGATGATGATGATGATGATGATGAT 1019
Db      1232 ACATTTTATTTAAATGATGATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291
Oy      1020 TGTAAAGCAAGCAAGATATAGTGTGTGTATTTAGAAATGCTTTTGGATTA 1079
Db      1292 TCTTAAAGCAATTAAGCACTATATAGCTATTAAGATCTGAAGAGGTGTCTGGGGTTA 1351
Oy      1080 A 1080
Db      1352 A 1352

```

```

RESULT 12
AB023628          1832 bp   DNA          linear   VRT 02-APR-1999
LOCUS            AB023628          1832 bp   DNA          linear   VRT 02-APR-1999
DEFINITION      Danio rerio 2FT2 gene for alpha(1,3)fucosyltransferase, complete
ACCESSION      AB023628
VERSION        AB023628.1 GI:4587297
KEYWORDS        alpha(1,3)fucosyltransferase.
SOURCE          Danio rerio DNA.

```

```

ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE      1 (sites)
AUTHORS        Kageyama, N., Natsuka, S. and Hase, S.
TITLE          Molecular cloning and characterization of two zebrafish
                alpha(1,3)fucosyltransferase genes developmentally regulated in
                embryogenesis
JOURNAL        J. Biochem. (1999) In press
REFERENCE      2 (bases 1 to 1832)
AUTHORS        Kageyama, N., Natsuka, S. and Hase, S.
TITLE          Direct Submission
JOURNAL        Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate
                School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034,
                Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, Tel: 81-6-6850-5381,
                Fax: 81-6-6850-5383)
FEATURES             Location/Qualifiers
     source           1..1832
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
     gene             332..1510
                     /gene="zft2"
     CDS              332..1510
                     /product="alpha(1,3)fucosyltransferase"
                     /protein_id="BA76707.1"
                     /db_xref="GI:4587298"
                     /translation="MDKLTVMKIKLTPSSSTOKLIVTFMLISLCIFVYVNPNT
                     TFFKPTMGHANSSEIETLHLMQVRECIIMKASYNALPTAPNPAPKADODIIVL
                     IMAFPRVFPDLKDLGLEYNTHHCOITVDRSLIQAHVMHNRMSDFDPKPPDF
                     OKIWMNNESEPLHNDLNDRENTVSRYNSDIPVYGRVLDATGEOKFTPK
                     DKIVCMVNSFGNHRHFFERFAKHINVSSTGGHFNMRNGDYGAVVSSCFEYLS
                     FENSIHRDYTEKLENPALCTGVPIVYPPRPNRYEPIPRPAFIHVDFPPKELADH
                     LISDQNEQYKQPFNMRKLYVSTSGLEHSCIKDYLKRNHYIAVTDLKQFWG"
BASE COUNT      585 a      394 c      317 g      536 t
ORIGIN
Query Match      21.1%; Score 228.2; DB 5; Length 1832;
Best Local Similarity 55.0%; Pred. No. 3.1e-50;
Matches 496; Conservative 0; Mismatches 393; Indels 12; Gaps 2;
Oy      183 TAAGCAATACCATCTCTGTTGGTATGCGCATTTGGCAGACCTTGAACCTTATC 242
Db      619 TCAGACATCATCGTGTATATCTGATGCTCTCTTCAAGATACCTTTGACCTGAAAGA 678
Oy      243 CTGCCAAGCAATGTTCATATATTCAGAGGTCCTCATTCACAGACAGCCGCTCATTTGACAA 302
Db      679 TTGTGTTTGGAGTTAAATCATCCAGCGCTGTCATTAATACATACAGAGAGCTGATCA 738
Oy      303 CAATTCACATGCGGCTCTATTCACCATAGACATCACTGGGATCTGATTAACCTTACC 362
Db      739 GAAAGCGCATGGAGTTATCTTCCACACAGATATGACACAGAT -----TTTCC 789
Oy      363 TCAGACAGCGCCAGCCACCTTTCAGAAATGATGATGATGATGATGATGATGATGATGAT 422
Db      790 GCAACCCCGCGACCAAGATTTCCAAAGTGATATGATGATGATGATGATGATGATGAT 849
Oy      423 CACCCCCCAAAAGAGTGGATGGAACACTTGTCAACCTGACTTAACCTTATGCGCGTGA 482
Db      850 TTGCATATCAAAATGACTTCTAAATGATGCTTTATGATGAGCTCAAGTTACAACAGGA 909
Oy      483 TTCAGATATCAAGTGCCTTATGCGCTTCTGACGCTGAGACAA-----AATCCTTTGTGT 539
Db      910 TTCAGATATTCAGTGCCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
Oy      540 TGAAGTGCACAGCAGAGAAAGTGTGTGCTGGGTTGCTGATTAAGTCAACCTGAGCA 599
Db      970 CACCAATCCCAAAAGAGATTAACCTGATGCTGATGATGATTAACCTTCAAGGCAACA 1029
Oy      600 TGCAGAGTCAAGTATTACAAAGAGCTCAGCAAGAGTATTTGAATTCACACCTATGGCA 659

```

	Db	1030	CAGGCGACACATTTCTTCTAGAGTTGGCCGAACACATTAAATGAGCAGCTTAATGAGG	1089
Oy	660	AGCATTCGGAGATATAGTGTAAGCATTAATAATCTGATTCGCCACATATGCTACTGTAAAT	719	
Db	1090	GCACTTCACAACATGCATTAATATGTCGACGATTAATGAAATGTCGTATCCAGTTGTAAT	1149	
Oy	720	TATCTCTTCATTTGAAAACCTCAATTCACAAAGATTACATCACAGAAAAGCTCTCAATGC	779	
Db	1150	CTACCGTCCTTTAGAACCTCCATTCACAGAGACTATTTCTATCGAAGAAGCTGTCAATTC	1209	
Oy	780	AATTTTGGCTGTCTCAGTACCTCTGGTCCCTGGGTCATCTYAGGAAAACTTATGAAATTA	839	
Db	1210	TTTTAGCAGCTGCTACAGTCTCTGCTAATTTAGGCCGCCCCAGATTAATCTATGAAAGATT	1269	
Oy	840	TATTCAGCTATTCATTCATTCATGTATGGAAGATTATTTATCTCCACAGTAGTAAGAA	899	
Db	1270	CATACCAGAGAGACCCCTTCATCCATGSGATGATTTCCCAATGCCAAGAACTGGCGGA	1329	
Oy	900	ATATCTGAAGGAGATGTACAAAACAAATTAAGTCTTACTTACTTACTTAATCTACGAA	959	
Db	1330	TCACCTTATTCCTTGAGCCANNAAGGAGAACATTAACAGATTTTTCACCTGAGAGAA	1389	
Oy	960	GGAATTTTACTGTAAACCTACCCAGGGTTTTGGGAATCATCATGCTGCTGACATCGACCA	1019	
Db	1390	ACTTTATGTTTCCAAGAGTACATCGTTGGGCTGGAACATTCCTGTAGAAATTTTGTGATTA	1449	
Oy	1020	TGTAAAAAGGCATCMAGATATATATAGTCTGTGTGATTAATTTAGAAATGTTTTGGAATTA	1079	
Db	1450	TCCTTAAACGATTAACACATATATACCGTTACGATCTGAAGAGGTGTTCTGCGGTTA	1509	
Oy	1080	A 1080		
Db	1510	A 1510		
RESULT 13				
LOCUS	AF288369			
DEFINITION	Gallus gallus alpha (1.3) fucosyltransferase gene, complete cds.			
ACCESSION	AF288369			
VERSION	AF288369.1	GI:17224489		
KEYWORDS	.			
SOURCE	Chicken.			
ORGANISM	Gallus gallus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
JOURNAL	Phasianinae; Gallus.			
TITLE	1 (bases 1 to 1264)			
REFERENCE	Foley,A.C., Schmitz,B., Stern,C.D. and Streifl,A.			
AUTHORS	Unpublished			
JOURNAL	2 (bases 1 to 1264)			
TITLE	Foley,A.C., Schmitz,B., Stern,C.D. and Streifl,A.			
AUTHORS	Direct Submission			
JOURNAL	Submitted (19-JUL-2000) Genetics and Development, Columbia			
TITLE	University, 701 West 168th Street, New York, NY 10032, USA			
JOURNAL	Location/Qualifiers			
FEATURES	1..1264			
source	/organism="Gallus gallus"			
	/db_xref="taxon:9031"			
CDS	<1..1062			
	/note="fuf"			
	/codon_start=1			
	/product="alpha (1.3) fucosyltransferase"			
	/protein_id="AI136994.1"			
	/db_xref="GI:17224480"			
	/translation="SMKLLTFLLSFTRFCLPASLRPRTPEPDVNSYSTERLP			
	AGTSALPSERKMILLTLMTWPGHFRFNSSNSELFDTPCHFTTRRSQKAADVIM			
	HHHDVDWDERLSRIPLRPSQRMIWFNMKSPPSHSNGLADHLFLNTSKYSRSDITL			
	PYEQLDLEDPSPSLPKPTKRWAVVWSMRSSHRYKYVEDEKKIIPVDYGRGHLEP			
	LSDMHIVDFASAGELAOYLDELISRPTEKYRYPQWRMLKLPIVLIGAMAYCRACHR			

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (sites)	Kaneke,M., Nishihara,S., Kitano,T., Natsumitsu,H. and Saitou,N.	The evolutionary history of glycosyltransferase genes	unpublished
2 (bases 1 to 1092)	Kaneke,M., Saitou,N. and Kitano,T.	Direct Submission	Submitted (17-DEC-1999) Mka Kaneke, National Institute of Genetics, Laboratory of Evolutionary Genetics; Yata Iili, Mishima, Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp, Tel:81-559-81-6790, Fax:81-559-81-6789)
FEATURES	SOURCE	LOCATION/Organism	QUALIFIERS
gene	1..1092	/organism="Xenopus laevis"	
CDS	/db_xref="taxon:8355"	/product="xFTI"	
	/issue_type="bone marrow"	/protein_id="BAB82490.1"	
	1..1092	/db_xref="GI:18146868"	
	/gene="xFTI"	/translation="MRQDSLIVSRILCMPLCRSRISLRSLAASLTCYLVVYCSODLAGPINFPLMGTTFPPKQVTVLYVEPGKKRRIGDCHMLFNVGCHITNRSLOEDAMFHRHIDSDPEDFRROSOKIMWNESPSPMVLASGIFNTWMSYVNDSDIFMPDYGFELFSKRRAKIIVLPKRLVMVVISNNMEDHERVOYAEINRYEI DYVGRGLDKEDNIYKTESEYEKFYLAEFSLHDVDYTEKLMMNAEKSNAPLYMGPS RYNEMPIPRSSFIHNDDSPSPRKMIANYIKLHKDNKHLYRYTTPWKRRIVHYHTSPD EMTCTAQSVKANGNRHMRTASDLAFGES"	
BASE COUNT	256 a	281 c	294 g 261 t
ORIGIN			
Query Match	19.1%; Score 206.6;	DB 5;	Length 1092;
Best Local Similarity	56.1%; Pred. No. 1.9e-44;		
Matches 474;	Conservative 0;	Mismatches 359;	Indels 12; Gaps 4;
OY	184 AACGAACATACCATTCTGGTTGGGTATAGGCATTGGGCGACGCTTGACTTACATCC	243	
Db	199 AAGCAAGTCACTGTCTTAATATGGTAGGCCATTGGCCAAAGCGCATTTGGGGAT	258	
OY	244 TGGCAGCAATGTTCAATATCCAAAGGTGCCATCTCACACAGACCGCTCATTTGACAC	303	
Db	259 TGCCACATGCTCTTCATATGTCACGGGGGCCACATCACACAAGAGAGCTTTACGAG	318	
OY	304 AAATCCCATGGGGTCGATTCACCATAGACATGACGCTGGGATCTGACAACTTAACCT	363	
Db	319 GAGCAGACGCCATTATGTTCCACCGGGACAT--TCGTATTCAGTAGACCTTCTT	375	
OY	364 CAGCAGGCCAGGCCACCTTTCAGAAATGATTTGGATGATTTAGAAGTCAACCATCAC	423	
Db	376 TTCCGAGAGGCGCCAGAGTCCCAAAGAGATCTGGATGAACTTCGMAATTCGCTCCAT	435	
OY	424 ACCCCCCAAAAGATGGCATGAACACTTGTTCACCTGACTCTAACCTTATGCGCGTAT	483	
Db	436 TCGCCCTGGGTGGGAGATCTGGGGGATCTTCAACATGAGACCATGTCCTACAGGATGGAC	495	
OY	484 TCAGATATCCAAGGCTTATAGGCTTCTGACGGGTAGGCAAAATCCCTTGTGTTGAA	543	
Db	496 TCGGATATCTTCAAGCCCTACAGGCTTCT--ATTCTCCAAAGAAACGTGCCAAATATGTC	552	
OY	544 GTGCCAAGCAAGAGAGATTGGTGTGGGTGTGGATGATACTGGAACCTGAGCATGCC	603	
Db	553 CTGGCGGCGCAAGAAAGAGCTGGTGGCTGGGTGATTCAGCAATCGGMAAGAGACAGGAA	612	
OY	604 AGGTCGAATGTTACAAAGAGCTCAGCAAGAGATTTGAATTCACACCTATGCGCAAGCA	663	
Db	613 AGGGTCCTGATTTCAACAGAGCTGAGGAATCTCGAAATCGAGCTTACGGCC---GG	669	
OY	664 TTCGGAAGATACGTAACGATATAAAATCTGATTTCCACCAATATCTACTGTGTAATTTAT	723	
Db	670 TATGGGCTGGATTTAAAAAGAGCAACATCGTCAAGACCGCTCGAGATATAAATTTTAC	729	

QY	724	CTTTCATTGGAAACCTCAATTCACCAACAAATATACATACAGAAAGACCTC---ACAAATGCA	780
Db	730	TTGGCTTTTGGAAATTCCTTGACACAGCATTTACATACACGGAAAGCTGTGGAGNACGCG	789
QY	781	TTTTTGGCTGGTTCAGTACCTGTGTTCCTGGGTCCATCTAGGAAAACTATGAGAAATTAT	840
Db	790	TTTTAAATCCAAAGCCATCCCATCGTCAATGGAGGCCACGCCCTATACTAGAGATGTTTC	849
QY	841	ATTCAGCTGATTCATTATTCATGTGGAAAGATTTAACCTCTCCAGTGAATTACGAAAA	900
Db	850	ATACGCGCAGCTCTTTTATTTCAGCTGACATATTTCTCCACCCAGGAAAGCTGGCCATG	909
QY	901	TATCTGAAGGAAGTTGACAAAAACAATATAGTTTGTACCTTAGTCTTTAACTGGAGAAAG	960
Db	910	TACTCGAAGCAGCTGGATTAATAAACAATTACTGTTACCCAGCATATTTTACTGTGAGAG	969
QY	961	GATTTTACTGTAAACCTTACCAAGGTTTTTGGGAAATCATGATGATCGCTGGACGACCAT	1020
Db	970	CGTTTACGACAGTGCAGCGTACGTCCTTTTGGGACGACGATTACTGACACCGCTGCCAAGC	1029
QY	1021	GTAAA 1025	
Db	1030	GTCAA 1034	
RESULT 15			
LOCUS	GGU73678	2657 bp	DNA
DEFINITION	Gallus gallus alpha-1,3-fucosyltransferase (CFP1) gene, complete cds.		
ACCESSION	U73678		
VERSION	U73678.1		
KEYWORDS	GI:1657998		
SOURCE	chicken.		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.		
AUTHORS	1 (bases 1 to 2657)		
TITLE	Lee, K.-P., Carlson, L.M., Woodcock, J.B., Ramachandra, N., Schultz, T.L., Davis, T.A., Lowe, J.B., Thompson, C.B. and Larsen, R.D. Molecular cloning and characterization of CFP1, a developmentally regulated avian alpha(1,3)-fucosyltransferase gene		
JOURNAL	J Biol. Chem. 271 (51), 32960-32967 (1996)		
MEDLINE	97115837		
REFERENCE	2 (bases 1 to 2657)		
AUTHORS	Lee, K.-P., Carlson, L.M. and Larsen, R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-OCT-1996) Immune Cell Biology Program, Naval Medical Research Institute, 8901 Wisconsin Ave., Bethesda, MD 20889, USA		
FEATURES	Location/Qualifiers		
source	1..2657		
gene	/organism="Gallus gallus"		
CDS	/db_xref="taxon:9031"		
	170..1240		
	/gene="CFP1"		
	170..1240		
	/gene="CFP1"		
	/note="similar to mammalian fucosyltransferase FucTIV"		
	/codon_start=1		
	/product="alpha-1,3-fucosyltransferase"		
	/protein_id="AAC60060.1"		
	/db_xref="GI:1657998"		
	/translation="MELGPRMSPARPCCRRRRRRMALLGALLGALLALALYCVREIR		
	RKSAAGRPEGVTVLLMDEPFRPWPAPACRRRYNTGGLSADRGRVGEAAVLFH		
	HRDLALYRGROGLPRGPDPQRPQRMWVNFSESHSLGLAGLFRNMVTSYRRSDG		
	FVYGLYIEPPSPRPFLVPPKRSRLVAVVINSNMDEHARVRYVYQLKELPIDYGARG		
	MALLEGSVTVSAKYKFLYAFENSOHDDYTEELTKRAFAASAAPVPLVIGRRANVRE		
	IPDSFLIHNDPSPRLATLYLKLKLEFDLKNKRSYRRTAMRKRYEVHTSFDEHYKVC		
	EAVRTAGNQLKTYONLAGTFES"		
BASE COUNT	581 a	712 c	818 g
ORIGIN			546 t

Query Match 16.2%; Score 175.2; DB 5; Length 2657;
Best Local Similarity 52.9%; Pred. No. 4.9e-36;
Matches 453; Conservative 0; Mismatches 388; Indels 15; Gaps 3;

QY 186 CGAAATCTATTCGTTGGTGGATGACATTTGGGCGACACCTTTGACCTTACATCCG 245
DB 334 CGAGGTGACGGTGTGTGTGGAGCCCTTCGGCCGCCCTGGGGCCCGGCGGACG 393

QY 246 CCAAGCAATGTTCAATATCCAAAGGGTCCATCCACACAGACCGCATTTGACACAA 305
DB 394 CCGAGAGCGCTACAAACATCAGCGGCTGCTCTAGAGCGCCGACCGGGGGGTACGGCA 453

QY 306 ATCCCATGCGGTCCTGATTCACCATAGACAT-----CAGCTGGGATCTGACTAA 356
DB 454 GGCTGGGGCGGTCTCTCCACACCGGACCTGGCGCTGACGCGCCGACGGGGCTGCC 513

QY 357 CTTACCTCAGCAGGCGCAGCCCTTTCAGAAATGGATTTGGATGATTTAGAGTCACC 416
DB 514 CCGCGGGCCCGCCCGCGACCCCGCGACGCGTGGGTGGATGATGATGACCTGACGCGCC 573

QY 417 CACTCACACCCCGCAAAAGAGTGGATGAACTGTTCACCTGACTTACTATATG 476
DB 574 CTGCGACATCCCGCGGGGCGTGGCGCGGCTCTTCACTGGACATGTCGTACCG 633

QY 477 CCGGTATTCAGATATCCAAATGCTTATGCTTCTTACGCGTGACGACACAAATCCCTTGT 536
DB 634 ACGGGACTCGGAGCTCTTCGTCCCTACGGGTACCTCTACGAGCCGCGTCCGCGGCGCC 693

QY 537 GTTTGAAGTCCCAAGCAGAGAGAGTGTGCTGCTGGTGTGAGTAACTGGAACCTTGA 596
DB 694 CTTCGTGCTGCTCGGAAAGCGGCTGTGGCTGTGGTGTATCAGCACTGGAACGAGGA 753

QY 597 GCATGCCAGGTCGAATTAACAGAGCTCAGCAGAGTATGAAATCCACACTATGG 656
DB 754 GCACGGCCCGGTGCTCTACTACCGCAGCTGACAGCAGCCTCCATCGACCTGTACGG 813

QY 657 CCAAGCATTCGAGATAGTGAATGAATAAATCTGATTTCCACATATCTTACTGTAA 716
DB 814 ---GGCGCGGGGATGGCGCTGTGGAGGCGAGCGGTGTGAAGCGGTGTGCGCTACAA 870

QY 717 ATTTATCTTTTCAATTTGAAACTCAATTCACAAAGATTACATGACAAAGCTCT--A 773
DB 871 GTTCTACCTGGCTTCGAGAACTCCACACACGAGACTACATCACGAGAAAGCTGTGAA 930

QY 774 CAATGATTTTGGCTGTGAGTACCTGTGTCTGGTCCATTAAGGAAACTATGA 833
DB 931 GAACGCTTCGCGCGCGACGCGGTGCTGTGTCTGTGGCCCGCGAGGGCCAACTACGA 990

QY 834 GAATTAATTCAGCTGATTCATTCATTCATGAGAAATTTTAACTCTCCAGTGAGTT 893
DB 991 GCGCTTCATCCCGCGCTCTTATCCACAGTGGAGACTTCCCGAGCCCAAGGCTGCT 1050

QY 894 AGCAAAATATCTGAAGAGTTGACAAAAACATAAATAGTGTATCTTACTTTACTG 953
DB 1051 GCGCACTACCTCTGAGTCTCTGATTAACAAACAGCCAGCTACAGAGGATTTGCGCTG 1110

QY 954 GAGAAAGATTTTACTGTAACCTACACGCTTTGGGATTCACATGCAATGCTGCGCATG 1013
DB 1111 GCGGAACAAGTATGAAGTCCAGTCACTCTTCTGGGATGAGCATTTACTGCAAGGTTTG 1170

QY 1014 CGACCATGTAAAAAG 1029
DB 1171 CGAGGCGGTGAGGACG 1186

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 121.198 Seconds
(without alignments)
15299,444 Million cell updates/sec

Title: US-09-744-748-3_COPY_115_1194

Perfect score: 1080
Sequence: 1 atgacatcaacatccaaag.....agaatggtttggaattaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1080	100.0	2036	21	AAZ92672	Murine alpha-1,3-f
2	1080	100.0	2170	21	AAZ92645	Murine alpha-1,3-f
3	931.2	86.2	2676	21	AAZ92647	Human alpha-1,3-fu
4	931.2	86.2	2822	21	AAZ92646	Human alpha-1,3-fu
5	129	11.9	1814	18	AAZ59505	Murine myeloid-1in
6	128.6	11.9	3594	18	AAZ85219	Mouse alpha-fucosy
7	128.2	11.9	1086	15	AAO56912	Lewis blood group
8	128.2	11.9	2042	18	AAZ80111	Glycosyltransferase
9	128.2	11.9	2043	12	AAO13530	Glycosyltransferase

10	128.2	11.9	2043	15	AAO56906	DNA encoding a gly
11	128.2	11.9	2043	18	AAZ76765	Human alpha1,3/4
12	128.2	11.9	2043	18	AAZ61675	Human alpha(1,3/1,
13	126.6	11.7	1316	15	AAO56911	DNA encoding a gly
14	126.6	11.7	1316	18	AAZ61679	Human alpha(1,3)-f
15	124.6	11.5	1654	15	AAO56905	pcDNA1-alpha-(1-3)
16	124.6	11.5	1654	18	AAZ61680	Human alpha(1,3)-f
17	123	11.4	1126	21	AAZ621135	Human low adenosin
18	123	11.4	1126	21	AAZ621137	Human low adenosin
19	123	11.4	6944	21	AAZ621137	Human low adenosin
20	123	11.4	6944	21	AAZ621137	Human low adenosin
21	122.8	11.4	1701	15	AAO77732	Human alpha-1,3-fu
22	122.8	11.4	1701	21	AAZ621136	Human alpha-1,3-fu
23	122.8	11.4	1701	21	AAZ621136	Human alpha-1,3-fu
24	120	11.1	1256	21	AAZ621134	Human low adenosin
25	120	11.1	1256	21	AAZ621134	Human low adenosin
26	120	11.1	1400	13	AAO31436	Human adenosine re
27	120	11.1	1488	15	AAO56910	Encodes a Hela cel
28	120	11.1	2175	12	AAO14382	DNA encoding a gly
29	120	11.1	2175	24	AAZ17082	Human cDNA clone 7
30	120	11.1	2861	12	AAO14383	Human cDNA clone 7
31	120	11.1	2861	21	AAZ21133	Clone 1 encoding 1
32	120	11.1	2861	21	AAZ21133	Human low adenosin
33	120	11.1	2861	21	AAZ21133	Human low adenosin
34	120	11.1	3647	12	AAZ17083	Human adenosine re
35	120	11.1	3647	12	AAZ17083	Human cDNA clone 1
36	120	11.1	3647	18	AAO56909	GDP-Fuc:beta-D-Gal
37	118.4	11.0	2175	11	AAZ61678	DNA encoding a gly
38	116.8	10.8	2175	11	AAO06691	Human alpha(1,3)-f
39	116	10.7	795	17	AAZ59506	CDX, a Mita involy
40	64.6	6.0	1578	23	AAZ13798	Human myeloid deri
41	64.6	6.0	4229	23	AAZ11867	Alpha-(1,3/1/4)-fu
42	53	4.9	1353	21	AAZ51685	Drosophila melanog
43	43.6	4.0	2198	21	AAZ51685	C. elegans alpha-1
44	42.4	3.9	14041	22	AAZ48024	Mung bean alpha-1,
45	40.8	3.8	398	22	AAZ64428	Internal control B
						Novel human polyu

ALIGNMENTS

RESULT 1	
AAZ92672	
ID	AAZ92672 standard; cDNA; 2036 BP.
AC	AAZ92672:
XX	
XX	05-JUN-2000 (first entry)
XX	
XX	
DE	Murine alpha-1,3-fucosyltransferase gene exon 1 DNA, SEQ ID NO:30.
XX	
XX	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW	brain; kidney; recombinant expression; transgenic animal; knockout
KW	animal; FUC-TV; drug screening; inhibitor; potentiators; diagnosis;
KW	treatment; cancer; murine; mouse; ss.
XX	
OS	Mus sp.
XX	
XX	
PN	MO200006708-A1.
XX	
PD	10-FEB-2000.
XX	
PF	29-JUL-1999; 99WO-JP04092.
XX	
XX	29-JUL-1998; 98JP-0213823.
PR	
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	
PI	Narimatsu H, Kudo T, Sasaki K;
XX	
DR	WPI: 2000-183120/16.
XX	
PT	Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x

PT epitope and an antibody recognizing it useful for diagnosis of brain
PR and kidney disease and cancer. -

Example 8; Page 162-164; 172pp; Japanese

CC The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-IV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA926645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA090955), and sequences AA926570-92673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).

SQ Sequence 2036 BP; 666 A; 365 C; 348 G; 657 T; 0 other;

Query Match	100.08;	Score 1080;	DB 21;	Length 2036;
-------------	---------	-------------	--------	--------------

```
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY	1	ATGACATCAATCCAAAGGCAATTCCTGGCCATTTCCTAATGCTGTGATCAATCCCTGGG	50
Dd	9	ATGACATCAATCCAAAGGCAATTCCTGGCCATTTCCTAATGCTGTGATCAATCCCTGGG	68
OY	61	TGCTTCATGCAATGTCGTGCTCAATTTATATCAAGCCACAAACAGTGGGCTTCAGTCCA	120
Dd	69	TGCTTCATGCAATGTCGTGCTCAATTTATATCAAGCCACAAACAGTGGGCTTCAGTCCA	128
OY	121	ATGGAGTCTGCAATCTCTGTGCTGAAATATGAAAAATTTCTTCCCAAAAAAGTATAT	180
Dd	129	ATGGAGTCTGCAATCTCTGTGCTGAAATATGAAAAATTTCTTCCCAAAAAAGTATAT	188
OY	181	TTTAAACGAATACCATCTTGTTGGTTGGATAGGGCATTTGGGGAGAGACTTTAGCCTTACA	240
Dd	189	TTTAAACGAATACCATCTTGTTGGTTGGATAGGGCATTTGGGGAGAGACTTTAGCCTTACA	248
OY	241	TGCTGCGCAAGCAATGTCAATATCCAAAGGATGCCATCTCAACAAGACCGCTCATTTGAC	300
Dd	249	TGCTGCGCAAGCAATGTCAATATCCAAAGGATGCCATCTCAACAAGACCGCTCATTTGAC	308
OY	301	AACAAATCCATGGGGTCTGATTCACCATAGAGACATCAGCTGGGATCTGACTAATCTTA	360
Dd	309	AACAAATCCATGGGGTCTGATTCACCATAGAGACATCAGCTGGGATCTGACTAATCTTA	368
OY	361	CCTCAGCAGGCGCAGCCACCCCTTCAGAAATGGATTTGGATGAATTTAGATGACACCCACT	420
Dd	369	CCTCAGCAGGCGCAGCCACCCCTTCAGAAATGGATTTGGATGAATTTAGATGACACCCACT	428
OY	421	CACACCCGCCAAAAGATGGGCAATGAGACATCTGTGTCAACCTGACTCAATCTATGCGCGT	480
Dd	429	CACACCCGCCAAAAGATGGGCAATGAGACATCTGTGTCAACCTGACTCAATCTATGCGCGT	488
OY	481	GATTCAGATATCCAAATGCTTATAGCTCTTGTGACGGTGAAGCAAAATCCCTTGTGTGTT	540

Accession	Gene	Species	Strain	Library	Insert Size (bp)	Read Length (bp)	Reads	Clones	Cloning Vector	Cloning Site	Cloning Efficiency (%)	Cloning Method	Cloning Conditions	Cloning Time (days)	Cloning Cost (\$)	Cloning Notes
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													
Db	489	GATTGAGATATCCAAAGTGGCTCTTATGGCTTCTTGAGGGTGAACCAATATCCCTTGTGTTT	548													

XX WPI; 2000-183120/16.
 DR P-PSDB; AAY80995.
 XX
 PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT epitope and an antibody recognizing it useful for diagnosis of brain
 PT and kidney disease and cancer.
 XX
 XX Claim 4; Page 127-134; 172pp; Japanese.
 PS
 CC The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA292645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA80995), and sequences AA292670-292673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).
 CC
 XX
 XX Sequence 2170 BP; 715 A; 395 C; 368 G; 692 T; 0 other;
 SQ
 Query Match 100.0%; Score 1080; DB 21; Length 2170;
 Best Local Similarity 100.0%; Pred. No. 2.3e-310;
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 CACACCCCCCAAGAGTGGCATTTGAACACTGTGTCACCTGACTTAATTCGCCGT 480
 DB 535 CACACCCCCCAAGAGTGGCATTTGAACACTGTGTCACCTGACTTAATTCGCCGT 594
 QY 481 GATTCAGATATCCAAATGGCTTATGGCTTCTTGACCGTGAGCAAAATCCCTTGTGTT 540
 DB 595 GATTCAGATATCCAAATGGCTTATGGCTTCTTGACCGTGAGCAAAATCCCTTGTGTT 654
 QY 541 GAAGTCCCAAGCAGAGAGAAAGTGGTGGCTGGCTGGTGAAGTGAACCTGAGCAT 600
 DB 655 GAAGTCCCAAGCAGAGAGAAAGTGGTGGCTGGCTGGTGAAGTGAACCTGAGCAT 714
 QY 601 GCCAGGCTCAATGATTATCAACAGAGCTCAGCAAGAGTGAATTCACACCTATGGCCAA 660
 DB 715 GCCAGGCTCAATGATTATCAACAGAGCTCAGCAAGAGTGAATTCACACCTATGGCCAA 774
 QY 661 GCATTCGGGAATACGCTGAACGATTAATTCATTCACCATATCTACTGTAAATTT 720
 DB 775 GCATTCGGGAATACGCTGAACGATTAATTCATTCACCATATCTACTGTAAATTT 834
 QY 721 TATCTTCAATTTGAAACATCAATTCACAAAGATTATCATCAGAAAGCTATACATGCA 780
 DB 835 TATCTTCAATTTGAAACATCAATTCACAAAGATTATCATCAGAAAGCTATACATGCA 894
 QY 781 TTTTGGCTGGTTCAGTACCTGTGTCTGGGTGCTATGAGGAAATATGAAATAT 840
 DB 895 TTTTGGCTGGTTCAGTACCTGTGTCTGGGTGCTATGAGGAAATATGAAATAT 954
 QY 841 ATTCCAGCTGATTCATTCATTCATGAGAGATTTTAACTTCCCGTAGTTAGCAAAA 900
 DB 955 ATTCCAGCTGATTCATTCATTCATGAGAGATTTTAACTTCCCGTAGTTAGCAAAA 1014
 QY 901 TATCTAAGAGAGTGTGACAAAACAATGATTTGATCTAGTACTTAACTGAGAGAAAG 960
 DB 1015 TATCTAAGAGAGTGTGACAAAACAATGATTTGATCTAGTACTTAACTGAGAGAAAG 1074
 QY 961 GATTTTACTGTAAACCTACACGGTTTGGGAATCATCATGCTGCGCATGCGACCAT 1020
 DB 1075 GATTTTACTGTAAACCTACACGGTTTGGGAATCATCATGCTGCGCATGCGACCAT 1134
 QY 1021 GTAAAGAGCATCAAGAAATATAGTCTGTGTAATTTAGAGAAATGTTTGGAAATTA 1080
 DB 1135 GTAAAGAGCATCAAGAAATATAGTCTGTGTAATTTAGAGAAATGTTTGGAAATTA 1194

RESULT 3
 AA292647
 ID AA292647 standard; cDNA; 2676 BP.
 XX
 AC AA292647;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.
 XX
 KW Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
 KW treatment; cancer; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 295..1374
 FT /tag= a
 FT /product= "Human alpha-1,3-fucosyltransferase"
 PN WO200006708-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-UP04092.
 XX

PR	29-JUL-1998;	98JP-0213823.
XX	(KYOW) KYOMA HAKKO KOGYO KK.	
PA	Narimatsu H., Kudo T., Sasaki K.	
PI	WPI: 2000-183120/16.	
DR	P-PDSB; AA80996.	
XX	Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x	
PT	epitope and an antibody recognizing it useful for diagnosis of brain	
PN	and kidney disease and cancer.	
PS	Claim 4: Page 143-150; 172pp: Japanese.	
XX	The invention relates to a novel alpha-1,3-fucosyltransferase which	
CC	transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine	
CC	(generating the Lewis x or y epitope). It does not transfer a fucose	
CC	moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine	
CC	and therefore does not generate the sialyl-Lewis x epitope. The	
CC	invention also relates to DNA sequences encoding alpha-1,3-	
CC	fucosyltransferase and expression vectors and host cells comprising	
CC	these DNA sequences. The invention additionally encompasses the	
CC	preparation of alpha-1,3-fucosyltransferase from a culture of	
CC	transformed cells or by expression of the protein in a transgenic animal;	
CC	antibodies which recognise alpha-1,3-fucosyltransferase; methods for	
CC	screening potential inhibitors or potentiators of	
CC	alpha-1,3-fucosyltransferase activity or expression; the preparation of	
CC	compounds having fucose-containing sugar chains by use of the protein;	
CC	and knockout non-human animals lacking alpha-1,3-fucosyltransferase.	
CC	Alpha-1,3-fucosyltransferase has a similar substrate range to the known	
CC	FUC-IV and is expressed mainly in brain and kidney tissues.	
CC	Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,	
CC	potentiators and inhibitors may be used for the treatment and diagnosis	
CC	of diseases of the brain and kidney, and of cancers. They may be used for	
CC	the identification of substances which affect the activity or expression	
CC	of alpha-1,3-fucosyltransferase; such substrates may be used	
CC	therapeutically. The knockout animals can be used to study the mechanisms	
CC	of action and expression of alpha-1,3-fucosyltransferase. Sequences	
CC	A4292646 and A4292647 represent cDNAs encoding human	
CC	alpha-1,3-fucosyltransferase (AA80996).	
XX		
SQ	Sequence 2676 BP; 819 A; 527 C; 476 G; 854 T; 0 other:	
	Query Match	86.2%; Score 931.2; DB 21; Length 2676;
	Best Local Similarity	91.4%; Pred. No. 4.3e+26;
	Matches 987; Conservative 0; Mismatches 93; Indels 0; Gaps 0	
OY	1 ATGACATCAACATCCCAAGGCATTCTGGCCATTTCCTAATCGTGCATCATCCTGGGC 60	
DB	295 ATGACATCAACATCCCAAGGAATTTCTGCCCATTTTAATGTCTGCATTATCCTGGGC 354	
OY	61 TGCTTATGSGCATGTCTGCTCATATTATATCAACGCCACCACAACACTGGGTCTTGAGTCA 120	
DB	355 TGTTCATGSGCATGTCTGCTCATATTATCAATCAACCAACCAACACTGGGTCTTGAGTCA 414	
OY	121 ATGAGCTCTGCAGTCTGTCTGCTGAATAAATGAAAAATTTCTCTGCACAAAATGATAT 180	
DB	415 ATGAGATTCAGCAGCTGTCTGCTGAATAAATGAAAAATCTCTTTCACAAAATGATAT 474	
OY	181 TTTAACCAATACCATTCGTGGTTGGGATATGGCCATTGGGCGAGACCTTGAACCTTACA 240	
DB	475 TTTAATGAATACATATCTGGGTGGGTGGGCGATTTGGGCGAGACCTTGAACCTTACA 534	
OY	241 TCCTGCCAAGCAATGTTCAATATCCAAGGGSCTCATCCAAACAGACGCTCATATGAC 300	
DB	535 TCCTGCCAAGCAATGTTCAACATCCAGAAGATCCCATCTCCACACGAGACCTTCACTGAC 594	
OY	301 AAAAATCCAGTGGGCTCGATTTACACATATAGACATGACCTGGGATCTGACTAATTA 360	
DB	595 AACAAATCCAGTGGGCTCTGATCTACATCCAGACATATCATGATGGGATCTGACAAATTTA 654	
OY	361 CCTCACAGGCCAGGCCACCCCTTCAGAAATGGATTTGGATGAATTTAGAGTCCACCACT 420	

[illegible]

Query Match	11.9%	Score 129	DB 18	Length 1814
Best Local Similarity	52.3%	Pred. No. 9,28-28		
Matches 335	Conservative 0	Mismatches 300	Indels 6	Gaps
08-MAY-1996; 96WO-US06427.				
PR 07-JUN-1995; 95US-0483151.				
XX (GEHO) GEN HOSPITAL CORP.				
PA Holgersson J, Seed B:				
PI MPI: 1997-108639/10.				
XX P-PSDB: AAM11820.				
XX New murine alpha-(1,3)-fucosyltransferase - for fucosylating an				
XX antibody to protect mammals against e.g. septic shock or septicemia				
XX				
PS Claim 6; Fig 3; 58pp: English.				
XX A cDNA clone (AAT59505) codes for murine myeloid lineage alpha-				
CC (1,3)-fucosyltransferase (AAM11820), an enzyme which has a relatively				
CC strict substrate requirement for sialylated N-acetylglucosamine,				
CC which can account for the presence of the sialyl-Le ^x epitope on				
CC murine cells, and which is more effective than Fuc-TIV in support				
CC of E-selectin-mediated COS cell adhesion. It was isolated from a				
CC murine myeloid 32D c33 cell cDNA library by identification of a				
CC clone capable of directing the expression of sialyl-Le ^x				
CC determinants. Transformed host cells (e.g. 32D c33 or human 293				
CC cells) expressing the alpha-(1,3)-fucosyltransferase, and pref.				
CC another fucosyltransferase such as human Fuc-TIV (see also AAT59506),				
CC can be used to fucosylate an antibody (e.g. IgG or IgM) for use in				
CC protecting an animal against an adverse immune reaction, esp. septic				
CC shock or septicemia.				
XX				
SQ Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other;				
Query Match	11.9%	Score 129	DB 18	Length 1814
Best Local Similarity	52.3%	Pred. No. 9,28-28		
Matches 335	Conservative 0	Mismatches 300	Indels 6	Gaps
269 GGTGCGCATCTCACACAGACCGCTCACTTGTACAAATAATCCGAGCGGTCTGATTCAC 328				
Db 548 GCTCCCTCTGATGCTTACCGGAGCGCTCTGACCGAGTCTGATGCTCTTCAC 607				
OY 329 ATGAGACATCAGCTGGGATCTGACTTAATCTACCTCAGCAGCGCCACCTTTGCA 388				
Db 608 ACCGTGAGCTGCAAAACCCGGCAATCTCTCTACCCCTGACACAGAGCCACACGAGAGC 667				
OY 389 AATGATTTGGATGAAATTAGAGTCCACCCACTCACACCCCCCAAAAGAGTGCAATTGAC 448				
Db 668 CTGCGGCTGGGCTCCATGAGGAATGCGCCAGTAATACCAATGCTCTCATTCGCTTCGGG 727				
OY 449 ACTTGTTCAACCTGACTCTAATCTTATGCGCGGATTCAGATATCCAGTCCCTTATGGCT 508				
Db 728 GCATCTTCAACTGGGTGCTGAGTCTATCGCGCGATTCAGATATCTTTGTAACCTCAGGTC 787				
OY 509 TCTTGACGCTGAGACACAAATCCCTTTGTTGTTGAAGTCCCAAGCAAGAAAGTGGTC 568				
Db 788 GCTTGAGAGCTCTCTCTGGGCC---ACATCCCTCACTACCGGCCMAAAGCAGATGGCT 844				
OY 569 GCTGGTGTGATTACTGGAACCTGAGACATGCCAGGTCAGATTAATCAAGAGTCA 628				
Db 845 CTTGGTGATCAGCAATTTCCAGAGCGGACACACGTCGAAACCTGTACCGGAGCTGG 904				
OY 629 GCAGAGATTTGAAATTCACACACTTATGCGCAAGCTTGGAGAAATACGTGAAGATTA 668				
Db 905 CCCCCTATCTGCAAGTGAATGTGTTCGTCGCGCCAGAGGAGAGGCCCTTATGGCTTA 964				
OY 689 ATTCGATATCCACCATATCTACTCTGTAATTTTATCTTCAATTGAAACTCAATTCACA 748				
Db 965 GTCTGCTGCCCACTTTGGCCCGGATACGCTTCTTACTCTG6CTTTGAGACTCAACAGATC 1024				
OY 749 AAGATTACATCAGAAAAGCTCTA---CAATGCAATTTTGGCTGTTCAGTACTCTG 805				
Db 1025 GGGCATCACTACCTGGAAGTCTTGGCCGAATGACCCCTGGCGGCTGCTGACCGTGG 1084				

Oy	806	TCCGGGTCACATCTAGGAGAAATATGAATAATTATATCCAGCGCATTCATTCATC	865
Dd	1085	CGGTGGACCTCCTCGGGCACACTACGAGGCTTTTGCCACACAATGCTTTGTACACG	1144
Oy	866	TGGAGATTTTAACCTCCCACTAGTAGCAAAATATCTG	906
Dd	1145	TGCACGACCTTCACGCTCTGCCCTGTAAGCTGCTCTTCCTG	1185
RESULT 6			
AAT85219			
ID	AAT85219	standard; DNA; 3594 BP.	
XX			
AC	AAT85219;		
XX			
DT	02-MAR-1998	(first entry)	
XX			
DE	Mouse alpha-fucosyltransferase Fuc-TVII gene.		
XX			
KW	Alpha-fucosyltransferase; Fuc-TVII gene; mouse;		
KW	sialyl Lewis x determinant; oligosaccharide; antiinflammatory;		
KX	Inhibitor; ds.		
XX			
OS	Mus musculus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	996..3082	
FT		//tag= g	
FT		/note= "includes introns"	
FT	exon	465..477	
FT		//tag= b	
FT		/number= 1	
FT	Intron	478..669	
FT		//tag= c	
FT	exon	670..1149	
FT		//tag= d	
FT		/number= 2	
FT	Intron	/codon.start= 996..998	
FT		1150..1946	
FT		//tag= e	
FT	exon	/number= 2	
FT		1947..1959	
FT		//tag= f	
FT		/number= 3	
FT	Intron	/codon.start= 1947..1949	
FT		/note= "alternative ATG start codon"	
FT		1960..2065	
FT		//tag= g	
FT		/number= 3	
FT	exon	2067..3552	
FT		//tag= i	
FT		/number= 4	
FT		/codon.start= 2126..2128	
FT		/note= "alternative ATG start codon"	
PN	WO9732869-A1.		
PD	12-SEP-1997.		
PP	07-MAR-1997;	97MO-US03184.	
PR	08-MAR-1996;	96US-0613098.	
PA	(UNMI) UNIV MICHIGAN.		
PI	Gersten KM, Lowe JB, Natsuka S;		
PJ	WPI; 1997-470535/43.		
DR	P-PSDB; AAM27138 and AAM26671.		
PT	Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g.		

RESULT 6	
ID	AAT85219
XX	AAT85219 standard; DNA; 3594 bp.
AC	AAT85219;
XX	
DT	02-MAR-1998 (first entry)
XX	
DE	Mouse alpha-fucosyltransferase Fuc-TVII gene.
XX	
KW	Alpha-fucosyltransferase; Fuc-TVII gene; mouse;
KW	glycyl Lewis x determinant; oligosaccharide; antiinflammatory;
KW	inhibitor; ds.
XX	
OS	Mus musculus.
XX	
XX	Key
EH	Location/Qualifiers
FT	CDS
FT	996..3082
FT	/*tag= a
FT	/note= "includes introns"
FT	465..477
FT	/*tag= b
FT	/number= 1
FT	478..669
FT	/*tag= c
FT	/number= 1
FT	670..1149
FT	/*tag= d
FT	/number= 2
FT	/codon_start= 996..998
FT	1150..1946
FT	/*tag= e
FT	/number= 2
FT	1947..1959
FT	/*tag= f
FT	/number= 3
FT	/codon_start= 1947..1949
FT	/note= "alternative ATG start codon"
FT	1950..2055
FT	/*tag= g
FT	/number= 3
FT	2057..3552
FT	/*tag= f
FT	/number= 4
FT	/codon_start= 2126..2128
FT	/note= "alternative ATG start codon"
XX	
PN	W09732889-A1.
XX	
PD	12-SEP-1997.
XX	
PF	07-MAR-1997; 97WO-US03184.
XX	
PR	08-MAR-1996; 96US-0613098.
XX	
PA	(UNMT) UNIV MICHIGAN.
XX	
PI	Gersten KM, Lowe JB, Natsuka S;
XX	
DR	WPI; 1997-470535/43.
DR	P-PSDB; AAW27138 and AAW26671.
XX	
PT	Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g

PT screening for inhibitors used as antiinflammatory agents
 XX Claim 1; Fig 2; 91pp; English.
 XX

CC This DNA sequence comprises the mouse Fuc-TVIII gene encoding
 CC alpha-fucosyltransferase. It was isolated from a 313 cell genomic
 CC library by screening with a probe from the catalytic domain of the
 CC human Lewis alpha(1,3/1,4)fucosyltransferase (Fuc-TVIII).
 CC Examination of the mouse Fuc-TVIII locus identified 3 Met codons
 CC that may initiate translation of alpha(1,3) fucosyltransferases
 CC with different cytosolic domains encoded by exons 2 and/or 3, but
 CC with identical Golgi-localised catalytic domains encoded by exon 4
 CC (see AA027138 and AA026671). The proteins and nucleic acids can
 CC be used to construct animal cell lines able to post-translationally
 CC modify oligosaccharides on cell surface, intracellular or secreted
 CC proteins and lipids, to isolate reagents for efficient enzymatic
 CC production of oligosaccharides, to generate antibodies to
 CC glycosyltransferases useful as diagnostic reagents, to screen for
 CC fucosyltransferase inhibitors and inactivators, especially those
 CC that act as antiinflammatory agents, for genotyping individuals at
 CC the fucosyltransferase locus and for in vitro synthesis of sialyl
 CC Lewis x tetrasaccharide.
 XX

SQ Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other:

Query Match 11.9%; Score 128.6; DB 18; Length 3594;
 Best Local Similarity 52.5%; Pred. No. 1.7e-27;

Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

DB 269 GGTGCATCTCACACACCGCTCATTTGACAAATCCAGCGGCTCTGATTACACC 328
 DB 2277 GCTGCGGTGTGAGTGTGATACCGAGCGCTGTACAGTGCTGATGCTGCTTCACC 2336
 OY 329 ATAGACATACACCTGAGTCTGATTAATCTACCTACGACGCGACCCCTTCAG 388
 DB 2337 ACCGTAGCTGCAACACCGGCAATCTCTCTACCCCTGGACAGAGCCACAGGACG 2396
 OY 389 AATGATTTGATGATGATTTAGATCAACCACTCACACCCCAAGAGTGCATTGAAC 448
 DB 2397 CTGGGCTGCGGCTCCATGAGATCCGCCAGTATACCATGATGTCATCGCTTCCGG 2456
 OY 449 ACTTGTCAACCTGACTTACTTATTCGCGGTGATTCAGATTCACAGTCCCTTAGGCT 508
 DB 2457 GCATCTCTCACTGCTGCTGAGTATGCTATGCTGATTCAGATTCCTTACCTAGGTC 2516
 OY 509 TCTTGACGGTGACCAAAATCCCTTGTGTGATGATGCAACGACGAGAAAGTGTGT 568
 DB 2517 GCTTGAAGCTCTCTCTGGGCC--ACATCCCACTACCGGCAAAAGCAGATGGCTG 2573
 OY 569 GCTGGTGTGATTAATGTAACCTGAGCATGCCAGGTCATGATTTACACAGACTCA 628
 DB 2574 CCGTGGTGATACAGCAATTCGAGGAGCGGAGCAGCGTGCAGAACTGTACCGGCACTG 2633
 OY 629 GCAAGAGTATGTAATCCACACCTATGCGCAAGCATTCGAGAAATACGTAAATAA 688
 DB 2634 CCCCTCATCTGCGAGGTGATGTTGCTGCGCCAGCGACGCCCTATGCGCTAAAT 2693
 OY 689 ATCTGATTCACCATATATCTACTTAAATTTATCTTTATCTTGAATCTCAATCACA 748
 DB 2694 GTTGTGCTGCCACTTTGGCCCGGTACCGCTTCTACTGCGCTTGGAACTCAACAGATC 2753
 OY 749 AAGATTACATCACGAAAGCTCTA--CAATGATTTTGGCTGCTGATCACTGTTG 805
 DB 2754 GGGACTACATCACTGAGAGTTCGCGCAATGCCCTGGGGGCTGTGTGCTATCCGCTG 2813
 OY 806 TCTTGGTGCATCTAGGAAACTATGAGATTAATTCAGCTGATTCATTCATTCAG 865
 DB 2814 CGGTGGACCTCTCTCGGCGCACCTACAGAGCTTTTGTGCACAGATGCTTTGTACAG 2873
 OY 866 TGAAGATTTTAACTCCAGTGAAGTAAAGC 896
 DB 2874 TGAACGACTTACGCTCTGCGCGTGAAGTGGC 2904

RESULT 7
 AA056912
 ID AA056912 standard; DNA; 1086 BP.
 AC
 AC AA056912;
 XX
 XX

DT 26-JUL-1994 (first entry)

DE Lewis blood group fucosyltransferase DNA.

XX Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
 KW surface; oligosaccharide; ss.

XX W09402616-A.

PD 03-FEB-1994.

PF 20-JUL-1993; 93WO-US06703.

PR 20-JUL-1992; 92US-0914281.

XX (UNMI) UNITV MICHIGAN.

XX Lowe JB;

DR WPI; 1994-048874/06.

XX DNA fragment encoding a glycosyltransferase - can be used for in
 PT vitro reactions to modify cell surface oligosaccharide(s) e.g.
 PT blood gp. determinants, to protect against transplant rejection

PS Disclosure; Fig 6; 249pp; English.

CC The sequence is that encoding a Lewis blood group fucosyl
 CC transferase (Fuc-TVIII). The enzyme produced by the DNA may be
 CC non glycosylated. This prevents premature loss of enzyme activity.
 CC It can also be used in in vitro reactions to modify cell surface
 CC oligosaccharide mols. e.g. blood group determinants.
 CC See also AA056905-11.

SQ Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;

Query Match 11.9%; Score 128.2; DB 15; Length 1086;
 Best Local Similarity 53.0%; Pred. No. 1.2e-27;

Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

OY 271 TGGCATCTCACACACCGCTCATTTGACAAATCCATGCGTCTGATTACCAT 330
 DB 271 TGGCATCTCACACACCGCTCATTTGACAAATCCATGCGTCTGATTACCAT 330
 OY 331 AAGAGATGACGCTGGGA---TCTGACTAATCTACTCAGCAGGCGGACCCCTTTCAG 387
 DB 331 TGGGATATATGTCACACCTTAATGTCAGCTCTCCACCTTCCCGGCGGAGGGCAG 390
 OY 388 AATGATTTGATGATGATTTAGAGTCAACCACTCACACCCCAAGAGTGGCATTTGA 447
 DB 391 CGGTGATCTGTTTCACTTGAAGCACCACCCCTTAATCTGCGACAGCAGCCTTGAC 450
 OY 448 CACTTGTCAACCTGACTCTAATCTTATGCGCGGTGATTCAGATATCCAGTGCCTATGAC 507
 DB 451 AGATATCTCAATCTCAATCTGTCATGTCACGAGAGTCTCGACATCTTACGCGCTACG 510
 OY 508 TTTCTACGCTGAGCAAAATCCCTTTG-----TGTGTAAGTGCACCAAGAGG 558
 DB 511 TGGCTGACCGGCTGTGCGGCCACGCTGCGCACCAACGCTCAACCTCTGCGCAAGACC 570
 OY 559 AAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
 DB 571 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
 OY 619 AACGAGCTACGCAAGATTTGAAATTCACACCTATGCGCAAGCATTTGAGAAATACG 678


```

XX AC AAQ13330;
XX XX
XX 07-NOV-1991 (first entry)
XX DE Glycosyltransferase gene in clone pCDM7-alpha(1,3/1,4)FT.
XX XX
XX KM ss.
XX OS Homo sapiens.
XX XX
XX PH Key Location/Qualifiers
XX FT CDS 73..1158
XX FT polyA_site /tag= a
XX FT polyA_site 2023..2034
XX FT polyA_signal /tag= b
XX FT repeat_region 1926..2003
XX FT repeat_region /tag= c
XX FT repeat_region 1183..1547
XX FT /tag= d
XX FT /note= "20 copies of a degenerate 16bp sequence"
XX FT repeat_region 1885..1999
XX FT /tag= e
XX FT /note= "similar to human Alu sequences"
XX FT
XX PN WO9112340-A.
XX XX
XX PD 22-AUG-1991.
XX XX
XX PF 14-FEB-1991; 91WO-US00899.
XX XX
XX PR 12-DEC-1990; 90US-0627621.
XX PR 14-FEB-1990; 90US-0479858.
XX PR 14-FEB-1990; 90US-0480133.
XX XX
XX PA (UNMI ) UNIV OF MICHIGAN.
XX XX
XX PI Lowe JB;
XX XX
XX DR WPI; 1991-267151/36.
XX DR P-PSDB; AAR13749.
XX XX
XX PT Isolation of gene conveying post-translational characteristic -
XX PT e.g. the presence of soluble or membrane bound oligo or
XX PT polysaccharide or glycosyltransferase.
XX XX
XX PS Claim 11; Fig 1; 155pp; English.
XX XX
XX CC The DNA encodes a protein sequence capable of functioning as a
XX CC GDP-Fuc: [beta-D-Gal(1,4/1,3)]-D-glucNAc(Glc)alpha (1,3/1,4)-
XX CC fucosyltransferase. The sequence coded by nucleotides 127 to 1083
XX CC represents the functional protein. The DNA sequence can be used to
XX CC construct animal cell lines with respect to posttranslational
XX CC modification of the oligosaccharides on cell-surface, intracellular
XX CC or secreted proteins or lipids by sub-terminal alpha(1,3) and alpha
XX CC (1,4) fucose residues that represent the products of this enzyme.
XX CC These can be used for the prodn. of diagnostics and therapeutics.
XX CC See also AAQ13331-Q13333.
XX XX
XX SQ Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other:

Query Match 11.9%; Score 128.2; DB 12; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1,7e-27;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

```

```

OY 388 AATGATTTGGATGATGATTTAGATGTCACCACTCACACCCCCCAAGAGTGGCATTTGAA 447
OY 463 CGCTGGATCTGGTTCACCTTGGAGCCACCCCTTAAGTGCACAGCCTGGAACCTTGAC 522
OY 448 CACTTGTTCACCTGACCTCTAATCTATGCGCGTGAATCAGATATCCAAAGTCTTATGCG 507
OY 523 AGATATCTTCACCACTCACCATTGCTTACCGACGACGACTCGACATCTTCAGCCCTACGCG 582
OY 508 TTCTTGACGGTGGAGCAACAAATCCCTTTG-----TGTGAGATGCGCAAGAGAG 558
OY 583 TGGCTGGAGCCCTGGTCCGGCCAGCCTGCCCCACCAACCGCTTAACCTTCCGCGCAAGACC 642
OY 559 AAGTTGGTGTGCTGGCTGTGTGAGTAACAGGAACCTGAGACATGCGAGGCTCAAGTATTAAC 618
OY 643 GAGCTGTGTGCTGGCGGTGTGCAACTGGAAGCCGGAATGCAAGGAGGTGGCTACTTAC 702
OY 619 AAGGAGCTCAGCAAGATTTGAATTCACACCTTGTGCGCAAGCAATTTGGAGATATGCTG 678
OY 703 CAGAGCTGCGAGGCTCAATCTCAAGGTGAGCGTGTACGAGACGCTCCACACA---AGCCCTG 759
OY 679 AACGATTAATAATCTGATTTCCACCATATCTACTGTGAATTTATCTTCAATTTGAAAC 738
OY 760 CCCAAGGGGACCATATGAGAGACGCTGTCCGGTACAAAGTTTACTTGCGCTTCCAGAAC 819
OY 739 TCAATTCACAAAGATTACATCACAGAAAAGCTCT---ACAATGCATTTTGTGCTGTTCA 795
OY 820 TCCTTGACACCCCGACTACATCACGAAAGCTGTGAGGAAGGCCCTGAGGCCCTGGGCC 879
OY 796 GTACCTGTGTCTGCTGGCTCATCTAGGAAAACCTATGAGATTAATTCAGCTATTTCA 855
OY 880 GTGCCGTGTGTCTGGGCCCGCAGAGCAACCACTGAGAGGTCTCTCCACCCGACCC 939
OY 856 GTCATTCATGTGGAAGATTTTAATCTCCAGTAGGTTAGCAAAATATCTGGAAGATT 915
OY 940 TTCATTCACAGTGGACGACTTCCAGAGCCCAAGAGACTGTGCGCGTACTCTGAGAGACTG 999
OY 916 GACAAACAAATAGTTGATCTTACCTTACTTAACTGAGAGAG 960
OY 1000 GACAAAGACACGCGCGTACTGAGCTACTTTCCTGTGGCGGAG 1044

RESULT 10
AAQ56906
ID AAQ56906 standard; DNA; 2043 BP.
XX
XX AC AAQ56906;
XX XX
XX DT 26-JUL-1994 (first entry)
XX XX
XX DE DNA encoding a glycosyltransferase.
XX XX
XX KM Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
XX KM surface; oligosaccharide; ss.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 72..1158
XX FT /tag= a

OY 388 AATGATTTGGATGATGATTTAGATGTCACCACTCACACCCCCCAAGAGTGGCATTTGAA 447
OY 463 CGCTGGATCTGGTTCACCTTGGAGCCACCCCTTAAGTGCACAGCCTGGAACCTTGAC 522
OY 448 CACTTGTTCACCTGACCTCTAATCTATGCGCGTGAATCAGATATCCAAAGTCTTATGCG 507
OY 523 AGATATCTTCACCACTCACCATTGCTTACCGACGACGACTCGACATCTTCAGCCCTACGCG 582
OY 508 TTCTTGACGGTGGAGCAACAAATCCCTTTG-----TGTGAGATGCGCAAGAGAG 558
OY 583 TGGCTGGAGCCCTGGTCCGGCCAGCCTGCCCCACCAACCGCTTAACCTTCCGCGCAAGACC 642
OY 559 AAGTTGGTGTGCTGGCTGTGTGAGTAACAGGAACCTGAGACATGCGAGGCTCAAGTATTAAC 618
OY 643 GAGCTGTGTGCTGGCGGTGTGCAACTGGAAGCCGGAATGCAAGGAGGTGGCTACTTAC 702
OY 619 AAGGAGCTCAGCAAGATTTGAATTCACACCTTGTGCGCAAGCAATTTGGAGATATGCTG 678
OY 703 CAGAGCTGCGAGGCTCAATCTCAAGGTGAGCGTGTACGAGACGCTCCACACA---AGCCCTG 759
OY 679 AACGATTAATAATCTGATTTCCACCATATCTACTGTGAATTTATCTTCAATTTGAAAC 738
OY 760 CCCAAGGGGACCATATGAGAGACGCTGTCCGGTACAAAGTTTACTTGCGCTTCCAGAAC 819
OY 739 TCAATTCACAAAGATTACATCACAGAAAAGCTCT---ACAATGCATTTTGTGCTGTTCA 795
OY 820 TCCTTGACACCCCGACTACATCACGAAAGCTGTGAGGAAGGCCCTGAGGCCCTGGGCC 879
OY 796 GTACCTGTGTCTGCTGGCTCATCTAGGAAAACCTATGAGATTAATTCAGCTATTTCA 855
OY 880 GTGCCGTGTGTCTGGGCCCGCAGAGCAACCACTGAGAGGTCTCTCCACCCGACCC 939
OY 856 GTCATTCATGTGGAAGATTTTAATCTCCAGTAGGTTAGCAAAATATCTGGAAGATT 915
OY 940 TTCATTCACAGTGGACGACTTCCAGAGCCCAAGAGACTGTGCGCGTACTCTGAGAGACTG 999
OY 916 GACAAACAAATAGTTGATCTTACCTTACTTAACTGAGAGAG 960
OY 1000 GACAAAGACACGCGCGTACTGAGCTACTTTCCTGTGGCGGAG 1044

WPI; 1994-048874/06.

```

DR P-PSDB; AAR45934.
 XX DNA fragment encoding a glycosyltransferase - can be used for in
 PT vitro reactions to modify cell surface oligosaccharide(s) e.g.
 PT blood gp. determinants, to protect against transplant rejection
 XX
 PS Disclosure; Fig 1; 249pp; English.
 XX
 CC The sequence is that encoding human glycosyl transferase. The enzyme
 CC produced by the DNA may be non glycosylated. This prevents premature
 CC loss of enzyme activity. It can also be used in in vitro reactions to
 CC modify cell surface oligosaccharide mols. e.g. blood group determinants.
 CC See also AAO56505-12.
 XX
 SQ Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other:
 Query Match 11.9%; Score 128.2; DB 15; Length 2043;
 Best Local Similarity 53.0%; Pred. No. 1.7e-27;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
 OY 271 TGCATCTCACACAGACGCTCATTTGACAAACAAATCCCATGGCTCTGATTACCAT 330
 DB 343 TGCCACATCACTGCGCCGCGCAGAGGTATCCACAGGAGACACGGTCATCGACACAC 402
 OY 331 AGACATCAAGCTGGGA---TCGACTAATCTTACCTCAGCAGGCGACGCCCTTTGAG 387
 DB 403 TGGGATATCATCTCCACCCCTTAAGTCAGCCTCCACCTTCCCGAGGCGCGAGGGGCG 462
 OY 388 AATGATTTGATGATTAATTAAGTCACCCACTCACACCCCCCAAAAGAGTGCATTGAA 447
 DB 463 CGCTGGAATCGGTCACTTGAAGCCACCCCTTAACCTGCAAGCCCTGAAAGCCCTGGAC 522
 OY 448 CACTTGTTCACCTGACTGACTTAATTCGCGGTGATTCAGATATCCAAAGTGCCTTATG 507
 DB 523 AGATATCTTCAATCTCACCATGTCCTACGACGACGACTCCGACATCTTCAAGCCCTT 582
 OY 508 TTCTTGAAGGAGGACCAAAATCCCTTGG-----TGTTGAAGTGGCCCAAGCAGAGAG 558
 DB 583 TGGCTGAGCGCTGTGCTGCGGCCAGGCTGCGCCACCCAGCTCAACCTCTCTGGCGAAG 642
 OY 559 AAGTTGCTGTGCTGGGTGTGTGATTAATCTGAAACCTTGAGCATGCGAGGCTCAAGTAT 618
 DB 643 GAGCTGGTGGCTGGGCGGTGTCACAGGAGCGGAGCTCAGCGAGGGTGGCTTACTAC 702
 OY 619 AAGAGCTCAGCAAGAGTATGAAATCCACACTATGCGCCCAAGCATTCGGAGAAATAG 678
 DB 703 CAGAGCCTGCAGGCTCATCTCAAGGTGAGCTGACGAGACCTCCACA---AGCCCTG 759
 OY 679 AAGCATAAAAATCTGATTCACCATATCTACTGTAATTTATCTTTTCAATTTGAAAA 738
 DB 760 CCCAAGGGGACCATGATGAGGAGCGGTGCCGTTCAAGTTCTACCTGCGCTTCGAGAAC 819
 OY 739 TCAATTCACAAAGATTACATCACAGAAAAGCTCT---ACAATGATTTTGGGCTGTGA 795
 DB 820 TCCATTCACCCCGACATCATCATCACCGAGAGAGCTGTGAGAGAACGCCCTGGAGCG 879
 OY 756 GTACCTGTGTCTCGGCTCATCTTAGGAAAACTATGGAATTTATTCAGCTGATTTA 855
 DB 880 GTGCCCGGTGGTGGGCCCGACAGAACCACTACGAGAGTTCCTGACCCCGAGGCGC 939
 OY 856 TTCAATCATGTGAGAGATTTTAATCTCTCCAGTAGAGTATGCAAAATATCTGAAGAA 915
 DB 940 TTCAATCATGTGAGAGATTTTAATCTCTCCAGTAGAGTATGCAAAATATCTGAAGAA 999
 OY 916 GACAAAAACAATAAGTTAGTACCTTAGTACTTAACTTAACTGAGAAAG 960
 DB 1000 GACAAAGACCAAGCCGCGTAACTGAGCTACTTGTGCGTGGGAG 1044

RESULT 11
 AAT76769
 ID AAT76769 standard; cDNA; 2043 BP.

AC AAT76769;
 XX
 DT 15-SEP-1997 (first entry)
 XX
 DE Human alpha 1,3/4 fucosyltransferase cDNA.
 XX
 KW GDP-L-fucose:beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase;
 KW human; alpha 1,2 FT; transgene; transgenic mouse; animal model;
 KW intestinal adhesion; Helicobacter pylori infection; stomach;
 KW small intestine; gut; epithelial cell; surface receptor;
 KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
 KW gastric adenocarcinoma; Lewis antigen; fucosylation; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 104..1201
 FT /*tag= a
 XX
 PN US5625124-A.
 XX
 PD 29-APR-1997.
 XX
 PF 11-JUL-1994; 94US-0273411.
 XX
 PR 11-JUL-1994; 94US-0273411.
 XX
 PA (UNIV) UNIV WASHINGTON.
 XX
 PI Falk P. Gordon Jr;
 XX
 DR WPI; 1997-258275/23.
 DR P-PSDB; AAM23806.
 PT Animal model for Helicobacter pylori infection - comprising
 PT transgenic mouse expressing human enzyme promoting intestinal
 PT adhesion
 XX
 PS Example 1; Columns 23-26; 24pp; English.
 CC
 CC A claimed transgenic mouse expresses, in its intestinal epithelial
 CC cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
 CC fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
 CC beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
 CC called alpha 1,3/4 FT). The enzyme is expressed under the
 CC control of a gut epithelial cell-specific promoter and Helicobacter
 CC pylori adheres to the transgenic cells. The transgenic mouse and
 CC intestinal epithelial cells from it are useful as models for screening
 CC compounds for the ability to inhibit adhesion of H. pylori to gut
 CC epithelial cells. The present sequence encodes human alpha 1,3/4 FT
 CC and was published by Kukowska-Latallo et al., Genes Dev., 4:
 CC 1288-1303 (1990).
 CC
 SQ Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other:
 Query Match 11.9%; Score 128.2; DB 16; Length 2043;
 Best Local Similarity 53.0%; Pred. No. 1.7e-27;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
 OY 271 TGCATCTCACACAGACGCTCATTTGACAAACAAATCCCATGGCTCTGATTACCAT 330
 DB 343 TGCCACATCACTGCGCCGCGCAGAGGTATCCACAGGAGACACGGTCATCGACACAC 402
 OY 331 AGACATCAAGCTGGGA---TCGACTAATCTTACCTCAGCAGGCGACGCCCTTTGAG 387
 DB 403 TGGGATATCATCTCCACCCCTTAAGTCAGCCTCCACCTTCCCGAGGCGCGAGGGGCG 462
 OY 388 AATGATTTGATGATTAATTAAGTCACCCACTCACACCCCCCAAAAGAGTGCATTGAA 447
 DB 463 CGCTGGAATCGGTCACTTGAAGCCACCCCTTAACCTGCAAGCCCTGAAAGCCCTGGAC 522
 OY 448 CACTTGTTCACCTGACTGACTTAATTCGCGGTGATTCAGATATCCAAAGTGCCTTATG 507

```
D 523 AGATCTTCAATCTCACCAATGCTTACCCGACGACTCCGACATCTTACGCGCTACGCG 582
Q 508 TTCTTGACGGTGAACACAAATCCCTTTG-----TGTTTGAAGTCCACAGCAAGAG 558
D 583 TGCGTGGAGCGGTGTGTCGGGCGACGCGTCCGACCCGACCCGCTCAACCTCTGGGCAAGAC 642
Q 559 AAGTGGTGTGCTGGGTGTGAGTAACTGGAACCTGAGCATGCCAGGTCAAGTATAC 618
D 643 GACCTGGTGGCTGGGCGGTGTCCAACTGGAGACCGGACTCAGCCAGGCTGCGCTACTAC 702
Q 619 AAGAGCTGACGAAGATTTAAATCCACACCTATGCGCCAGCATTCGAGATACGTG 678
D 703 CACAGCGCTGACGCTCATCTCAAGGTGAGCTGACGAGCGTCCACACA--AGCCCTGTG 759
Q 679 AAGCATAAATATCGATCTCCACCATATCTACTTGTAAATTTATCTTCAATTTGAAG 738
D 760 CCAAGGGGACCATGATGAGAGCGCTGTCCGGTACAAAGTTCACCTGGGCTTCGAGAAC 819
Q 739 TCAATTCACAAAGATTTACATCAGAGAAAGCTCT--ACAATGCATTTTGGCTGTTC 795
D 820 TCCCTTGACCCCGACATCAATCAGCCGAGAAAGCTGTGAGAGAACGCCCTGGAGCGCTG 879
Q 796 GTACCTGTGTCTCGGGTCCATCTAGGGAAGAACTATGAGAAATATATTCAGCTGATTC 855
D 880 GTGCCGTGTGTGCTGGGCGCCACAGAGACGACCTAGAGAGTTCCTGCCACCCGAGCGC 939
Q 856 TTCAATTCATGAGAGATTTTAACTCTCCAGTGAAGTTAGCAAAATATCTGAAGAGATT 915
D 940 TTCAATCCAGCTGGAGCACTTCCAGAGAGCCCAAGAGACCTGGCCGGGTACCTGAGAG 999
Q 916 GACAAAAACAATTAAGTTGCTTACTTACTTTACTGGAAGAG 960
D 1000 GACAAAGACCAACGCCGCTACCTGAGCTACTTTCGTGGCGGAG 1044

RESULT 12
AAT61675
ID AAT61675 standard; cDNA; 2043 BP.
XX
AC AAT61675;
XX
DT 19-JUN-1997 (first entry)
XX
DE Human alpha(1.3/1.4)-fucosyltransferase (Lewis enzyme) cDNA.
XX
KM Alpha(1.3/1.4)-fucosyltransferase; Lewis enzyme; Fuc-TIII;
  glycosylation; oligosaccharide; blood group; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 73..1158
FT /tag= a
XX
PN MO9709421-A1.
XX
PD 13-MAR-1997.
XX
PF 06-SEP-1996; 96MO-US13816.
XX
PR 08-SEP-1995; 95US-0525058.
XX
PA (UNMI ) UNIV MICHIGAN.
XX
PI Legault DJ, Lowe JB;
XX
DR WPI; 1997-192897/17.
XX
P-PSDB; AAW13638.
XX
PT New recombinant fucosyltransferase proteins - useful for modifying
  cell surface oligosaccharide structures
XX
PS Example 3; Page 268-269; 329pp; English.
```

```
XX
CC A cDNA clone (AAT61675) codes for an enzyme (AAW13638) that functions as
CC a GDP-Fuc:(beta-D-Gal(1,4/1,3))-D-GlcNAc/(Glc) alpha(1,3/1,4)-
CC fucosyltransferase (Fuc-TIII) which, when expressed by the cloned
CC cDNA sequence, functions within mammalian cells to generate de novo
CC expression of specific SSEA-1 (Lewis x), sialyl-Lewis x, Lewis a and
CC sialyl-Lewis structures. The clone was identified by transfecting
CC COS-1 cells with genomic DNA from human A431 cells, and isolating a
CC cloned cDNA that detd. expression of surface-localised SSEA-1
CC structures. The cDNA can be used to construct animal cell lines
CC with specific capabilities with respect to post-translational
CC modification of the oligosaccharides of expressed proteins or
CC lipids, to produce recombinant Fuc-TIII for use in oligosaccharide
CC prodn., to genotype individuals for the Lewis locus, and in the
CC prodn. of chimeric fucosyltransferases (see also AAW14512-32).
XX
SQ Sequence 2043 BP; 413 A; 646 C; 576 G; 408 T; 0 other;

Query Match 11.9%; Score 128.2; DB 18; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1.7e-27;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

Q 271 TGGCATCTCACAACAGCCCTCATTTGTACAAACATCCATCGCTCCGATTCACCAT 330
D 343 TGGCATCTCACAACAGCCCTCATTTGTACAAACATCCATCGCTCCGATTCACCAT 402
Q 331 AGAGACATCAGCTGGGA---TGTACTAATCTTACCTCAGAGGCGCAGCCACCTTTAG 387
D 403 TGGGATATATATGTCACCAACCTTAAGTACAGCGCTCCACCTTCCCGAGCGCGAGGGGAG 452
Q 388 AAATGATTTGGATGATTTAGAGTCACCCGACATCACACCCCGCAAAAGAGTGGCATTTGA 447
D 463 CGCTGATCTGTTCAACTTGGAGCACACCCCTAATGTCAGACACCTGGAAGCCCTGGAG 522
Q 448 CACTTGTTCACCTGACTCTAATCTATCGCCGTGATTCAGATTTCCAGTGCCTTTAGC 507
D 523 AGATACTCTCAATCTACCAAGTCTACCGGACGACCTCCGACATCTTCAAGCCCTTACGCG 582
Q 508 TTCTTGACGGTGAACACAAATCCCTTTG-----TGTTTGAAGTCCACAGCAAGAG 558
D 583 TGCGTGGAGCGGTGTGTCGGGCGACGCTGCCACCCACCGCTCAACCTCTGGCGCAAGAGC 642
Q 559 AAGTGGTGTGCTGGGTGTGAGTAACTGGAACCTGAGCATGCCAGGTCAAGTATAC 618
D 643 GAGCTGGTGGCTGGGCGGTGTCCAACTGGAAGCCGAGACTCAGCCAGGCTGCGCTACTAC 702
Q 619 AAGGAGCTCAGCAAGATATTGAATTCACACCTATAGGCCAAGCATTCGAGATACGTG 678
D 703 CAGAGCCTGACGCTCATCTCAAGGTGAGCTGACGAGCGCTCCACACA--AGCCCTGTG 759
Q 679 AAGCATAAATATCGATCTCCACCATATCTACTTGTAAATTTATCTTCAATTTGAAG 738
D 760 CCAAGGGGACCATGATGAGAGCGCTGTCCGGTACAAAGTTCACCTGGCTTCGAGAAC 819
Q 739 TCAATTCACAAAGATTTACATCAGAGAAAGCTCT--ACAATGCATTTTGGCTGTTC 795
D 820 TCCCTTGACCCCGACATCAATCAGAGAGCTGTGAGAGAACCCCTGGAGGCTGGGCC 879
Q 796 GTACCTGTGTCTCGGTGCTCATCTAGGAAAAACTATGAGCAATTAATTCAGCTGATTC 855
D 880 GTGCCGTGTGTGCTGGGCGCCACAGAAAGCAATAGAGAGGTCTCCGACCCGAGCGCC 939
Q 856 TTCAATTCATGAGAGATTTTAACTCTCCAGAGAGTATGCAAAATATCTGAAGAGATT 915
D 940 TTCAATCCAGCTGGAGCACTTCCAGAGAGCCCAAGAGACCTGGCCGGTACCTGAGAG 999
Q 916 GACAAAAACAATTAAGTTGCTTACTTACTTTACTGGAAGAG 960
D 1000 GACAAAGACCAACGCCGCTACCTGAGCTACTTTCGTGGCGGAG 1044

RESULT 13
AA056911
```


SO Sequence 1316 BP; 236 A; 456 C; 372 G; 252 T; 0 other;

Query Match 11.7%; Score 126.6; DB 18; Length 1316;
Best Local Similarity 52.9%; Pred. No. 4e-27;
Matches 373; Conservative 0; Mismatches 314; Indels 18; Gaps 4;

```
QY 271 TGCATCTCACACACCGCTCTATTGTACACAAATCCATCGGCTGATTACCAT 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 TGCACATCTCACACCGCTCTATTGTACACAAATCCATCGGCTGATTACCAT 489
QY 331 AGAGACATCTCACACCGCTCTATTGTACACAAATCCATCGGCTGATTACCAT 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 TGCATCTCACACACCGCTCTATTGTACACAAATCCATCGGCTGATTACCAT 549
QY 388 AAATGATTTGATGATTTAGATGATGATGATGATGATGATGATGATGATGAT 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
QY 448 CACTTGTTCACCTGATCTTATCTTATGCTGATGATGATGATGATGATGATG 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 GGAATCTTCAATCTCAATCTTATCTTATGCTGATGATGATGATGATGATGAT 669
QY 508 TTTCTGACGGTGAACCAAAATCCCTTTG-----TGTGAGTGGCAAGCAAGAG 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 TGGCTGAGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 559 AAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789
QY 619 AAGGAGCTCAGCAAGATTTGAAATCCACCTATGCTGCTGCTGCTGCTGCTGCT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 790 CAGAGCTGCAAGCTATCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846
QY 679 AAGGATTTAAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 847 CCAAGGGGACCATGATGAGAGCTGTCCCGTACAAATCTTATGCTGCTGCTGCTG 906
QY 739 TCAATTCACAAAGATTTACATCAGCAAAAGCTCT---ACAATGATTTTGGCTGCT 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 907 TCCGTGACCCCGACATCATACAGCAAGATCTGTGAGGAGAGCCCTGAGGCTG 966
QY 796 GTACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 967 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
QY 856 TTTATTCATGATGATGATTTTAACTCTCCAGATGATGATGATGATGATGATGAT 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1027 TTTATTCATGATGATGATTTTAACTCTCCAGATGATGATGATGATGATGATG 1086
QY 916 GACAAAAACAATAAGTTTACTTACTTACTTACTTACTTACTTACTTACTTACT 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1087 GACAAAGAGACCGCCGCTACCTGAGTACTTTCCGCTGGGGGAG 1131
```

RESULT 15
AA056905
ID AA056905 standard; DNA; 1654 BP.
XX

AA056905:

26-JUL-1994 (first entry)

pcDNA1-alpha-(1-3)-Fuc-TVI genomic DNA.

XX Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;

XX surface; oligosaccharide; ss.

XX Homo sapiens.

XX key location/Qualifiers
FH 128..1208
FT CDS /tag- a

XX MO9402616-A.

XX 03-FEB-1994.

XX 20-JUL-1993; 93WO-US06703.

XX 20-JUL-1992; 92US-0914281.

XX (UNMI) UNIV MICHIGAN.

XX Lowe JB:

XX MPI; 1994-04874/06.

XX P-PSDB; AAR45933.

XX DNA fragment encoding a glycosyltransferase - can be used for in

XX vitro reactions to modify cell surface oligosaccharide(s) e.g.

XX blood gp. determinants, to protect against transplant rejection

XX Claim 3; Fig 7; 249pp; English.

XX The sequence is that encoding human GDP-Fuc: [beta-D-Gal (1-4)]-D-

XX GlucNAc alpha (1-3)-fucosyltransferase. The enzyme produced by the

XX DNA may be non glycosylated. This prevents premature loss of enzyme

XX activity. It can also be used in in vitro reactions to modify cell

XX surface oligosaccharide mols. e.g. blood group determinants.

XX See also AA056906-12.

SO Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match 11.5%; Score 124.6; DB 15; Length 1654;
Best Local Similarity 52.1%; Pred. No. 1.8e-26;
Matches 408; Conservative 0; Mismatches 354; Indels 21; Gaps 5;

```
QY 196 ATTCTGTTGGTATGCGCATTTGGCAGACCTTGTACCTTACATCTGCGCAAGATG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ATCTGCTGTGAGAGCTGCTTTTACAAACCCATACCTGTGCCCCCTGCTGAGATG 377
QY 256 TTTCAAT---ATCCAGGGGTGCGCATTCACAAAGACCGCTCATTTGTCAAAATCCAT 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTGCTGCGACGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
QY 313 GCGGTCTCTGATTCACATACATGATGATGATGATGATGATGATGATGATGATG 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GCGGTCTCTGATTCACATACATGATGATGATGATGATGATGATGATGATGATG 497
QY 370 GCCAGGCGACCTTTTGAAGATGATGATGATGATGATGATGATGATGATGATG 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CCGAGGGGCGAGGGGCGAGCATGATGATGATGATGATGATGATGATGATGATG 557
QY 430 CAAAGAGTGGGATGTAACACTTGTTCACACTGCTTACTTATGCGCGTGAATGAT 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CAGCTGAAGCCATGCGATGATGATGATGATGATGATGATGATGATGATGATG 617
QY 541 GAACTGCCAAGCAAGAAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AACCTCTCGGCGCAAGACCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
QY 601 GCCAGGTCAGATTTATGCAAGAGCTGAGCAAGATGTAATAATCCACACTATGCGCA 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCCAGGTCAGATTTATGCAAGAGCTGAGCAAGATGTAATAATCCACACTATGCG 794
QY 661 GCATTCGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CGCTCCCAAGAGCCCGCCAGGAGAACCATGATGAGAGCGCTGCTGCTGCTGCTGCT 854
QY 721 TATCTTCAATTTGAAACTCAATTCACAAAGATTTACACAGAAAGCTCTAC---AAT 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Tue Oct 8 10:18:19 2002

us-09-744-748-3_copy_115_1194.png

Page 14

Db	855	TATGTGGCCTTCGAGAACTCCTTCGACCCCGACTACATACCGAAGAACTGTGGAGGAAAC	914
Qy	778	GCATTTTGGCTGTGTTCTGACTGTGTTCTGTCCTGGCTCATCTAGGAGAAACTGATGAAAT	837
Db	915	GCCCTGGAGGCGTGGGCCGTGGCCGGTGGTGGGGCCCGACGAGAAGCAACTACAGAGAG	974
Qy	838	TATATTCACAGCTGATTCATTCATTCATGTGGAAGATTTTAATCTCCCGAGTGTACGA	897
Db	975	TTCTGCGCAACCGACGCGCTTCATCCACGTGGAGAGATTCCTCAAGGCCCAAGAGACTGGCC	1034
Qy	898	AAATATCTGAAGGAAGTGTGCACAAAACATAGTGTACCTTAGTTACTTTTAACTGTGAGA	957
Db	1035	CGGTACTCGAGGAGGCTGGACAGAAGACACGCGCGCTACTGAGTCTTGCTGGCGGG	1094
Qy	958	AAG	960
Db	1095	GAG	1097

Search completed: October 5, 2002, 23:10:19
Job time : 136.198 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 904.255 Seconds

(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-3_COPY_115_1194

Perfect score: 1080
Sequence: 1 atgacataacatccaagg.....agaaatggttggaataa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estlum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_lum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	661.8	61.3	697	9	BB397507 BB397507
2	580.8	53.8	607	10	B1102760 602888338
3	368	34.1	649	9	AV338075 AV338075
4	366	33.9	716	9	AV725071 AV725071
5	335.8	31.1	648	9	BB595164 BB595164
6	315	29.2	625	9	BB649967 BB649967
7	304.8	28.2	1101	12	CNS05957
8	303.8	28.1	767	9	BB640911 BB640911
9	272.2	25.2	579	10	BM021081
10	255.8	23.7	628	9	BB651409
11	203.4	19.4	568	10	BM021317
12	179.6	16.6	1100	12	CNS05792
13	175.8	16.3	844	12	CNS01V8V
14	171.4	15.9	844	12	BM082484
15	164.8	15.3	887	12	CNS02005
16	157.4	14.6	670	9	BB626126
17	156.8	14.5	1016	12	CNS0429P

18	152.2	14.1	199	9	BB398103
19	148.4	13.7	666	10	B050300
20	145.8	13.5	651	10	B036855
21	142	13.1	640	10	B073227
22	140.4	13.0	711	12	CNS034XF
23	133.4	12.4	501	12	B41073
24	118.6	11.0	674	10	BB605503
25	115.6	10.7	971	12	CNS0471V
26	115	10.6	533	10	B1604066
27	110.8	10.3	1023	12	CNS02ANG
28	110.2	10.2	615	10	B1326811
29	106.2	9.8	487	10	BB63396
30	106.2	9.8	720	10	BF180380
31	105.8	9.8	702	10	B0043578
32	90	8.3	863	12	CNS02ROF
33	88.4	8.2	551	10	B035701
34	87.8	8.1	566	10	B7090565
35	85.4	7.9	601	10	B036679
36	85.4	7.9	605	10	B028804
37	84.8	7.9	776	12	CNS03RIO
38	83.8	7.8	1056	12	CNS01VL7
39	82.6	7.6	638	10	B035697
40	81.8	7.5	655	10	B538740
41	81.4	7.5	878	12	CNS02Y1D
42	80.8	7.5	304	9	A0081478
43	80.8	7.5	1057	12	CNS03ADC
44	74.8	6.9	838	10	BG432556
45	74.6	6.9	590	10	BM082755

ALIGNMENTS

RESULT 1
LOCUS BB397507 697 bp mRNA linear EST 24-OCT-2001
DEFINITION BB397507 RIKEN full-length enriched, ES cells Mus musculus CDNA clone C330004C19 3' similar to AB015426 Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, mRNA sequence.

BB397507 2 GI:16411808

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE

AUTHORS

TITLE JOURNAL
COMMENT
Unpublished (2001)
On Jul 15, 2000 this sequence version replaced gi:9216903.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalisation and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura

FEATURES

BASE COUNT
ORIGIN

204

OY	404	ATTATGAGTCAACCCACATCAACACCCCAAAAGAGTGGCATTAAGCACTTTGTCAACCTGA	463
Db	4	AATTAGAGTCAACCCACATCAACACCCCAAAAGAGTGGCA-TGAACACTTTGTCAACCTGA	62
OY	464	CTCTAACTATACCCCTGATTTACATATCCAAATGCGTTATGGCTTTTGACGGTGAACA	523
Db	63	CTCTAACTATACCCCTGATTTACATATCCAAATGCGTTATGGCTTTTGACGGTGAACA	1233
OY	524	CAAAATCCCTTTGTGTTTGAAGGCCAAGCAAGTATGGTGTGGTGGTTGAGTA	583
Db	123	CAAAATCCCTTTGTGTTTGAAGGCCAAGCAAGTATGGTGTGGTGTGAGTA	1023
OY	584	ACTGGAACCTGAGCAATGCGAAGGTCAAGTATTACAAAGACATCAGCAAGTATTGAAA	643
Db	183	ACTGGAACCTGAGCAATGCGAAGGTCAAGTATTACAAAGACATCAGCAAGTATTGAAA	2424
OY	644	TCGACACTATGGCCAAAGCATTCGGAAGATTAAGTGAAGAGTAAAAATCTGATTTCCACCA	703
Db	243	TCGACACTATGGCCAAAGCATTCGGAAGATTAAGTGAAGAGTAAAAATCTGATTTCCACCA	3023
OY	704	TATCTACTGTAATATTTATCTTTCATTTGAAAACTCAATTCACAAAGATTACATCACAG	763

RESULT 2				
BII02760				
LOCUS				
DEFINITION	BII02760	607 bp	mRNA	linear EST 26-JUN-2001
ACCESSION	602888338F1	NCL_GAP_Kid14	Mus musculus	cDNA clone IMAGE:5045609
VERSION	BII02760			
KEYWORDS	5' , mRNA sequence.			
SOURCE	BII02760.1	GI:14553653		
ORGANISM	EST .			
	house mouse.			
	Mus musculus			
REFERENCE	Eumetazoa; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (Bases 1 to 607)			
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			

Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LMU at: <http://image.llnl.gov>
 Plate: LMU11119 row: 1 column: 10
 High quality sequence stop: 606.

FEATURES	
SOURCE	Location/Qualifiers
	1. 607
	/organism="Mus musculus"
	/strain="FVB/N"
	/db_xref="taxon:10090"
	/clone_image="5043609"
	/clone_lib="NCI CGAP Kidid4"
	/lab_host="B10B (T1 phage-resistant)"
	/note="Organ: kidney; Vector: pCMV-SORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. "
BASE COUNT	179 a 149 c 121 g 157 t 1 others
ORIGIN	
Query Match	53.88; Score 580.8; DB 10; Length 607;
Best Local Similarity	99.28; Pred. No. 2.4e-150;
Matches 604; Conservative	0; Mismatches 3; Indels 2; Gaps 2;

US-09-744-748-3_COPY-115_1194.rst

DB 2 ATCACTGTATTTATTTCTTTCTATTTGTAAGAACTCAATTCACAGATTAAGTACAGAA 61
 Oy 765 AAGGTCTACAAATCCATTTTGGCGGTGGTGTAGTACCTGTGTCTGGCTTCATTCAGAG 61
 Db 62 AAGGCTTACATATGCAATTTTGGGTGGTGTAGTACCTGTGTCTGGCTTCATTCAGAG 61
 Oy 825 AACTATGAAATTAATTTTCCAGTGTATTCATTCATTCATTCATTCATTCATTCATTCAG 824
 Db 122 AACTATGAAATTAATTTTCCAGTGTATTCATTCATTCATTCATTCATTCATTCATTCAG 824
 Oy 885 CAGTGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAG 884
 Db 182 CAGTGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAG 884
 Oy 945 CTCTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAG 944
 Db 242 CTCTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAG 944
 Oy 1005 CCGGCAATGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAG 1004
 Db 302 CCGGCAATGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAG 301
 Oy 1065 ATGCTTTGATTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAG 1064
 Db 362 ATGCTTTGATTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAGACAAAGCTTTTAACTGAG 361

RESULT 4
 AV725071
 LOCUS
 DEFINITION
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AV725071
 AV725071 HNC Homo sapiens
 AV725071.1 GI:10830091
 EST.
 human.
 Homo sapiens
 Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
 Primates; Catarrhini; Hominoidea; Homo.
 Lu, N., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
 Chen, J., Chen, Z., and Han, Z., Fu, G., Ren, S., Zhong, N., Lu, G., Hu, R.,
 Homo sapiens cDNA HNC clones
 Unpublished
 Contact: Zepeng Han
 Chinese National
 2012 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanz@cpic.sh.cn

TITLE
 JOURNAL
 COMMENT

FEATURES
 source

location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="HNCB1B02"
 /clone_1kb
 /tissue_type="HNC"
 /dev_stage="Hypothalamus"
 /lab_host="SOLiD"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT
 ORIGIN

219 a 113 c 133 g 249 t 2 others

Query Match
 Best Local Similarity 91.7%; Score 366; DB 9; Length 716;
 Matches 387; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Oy 659 AAGCATTCGAGATACGTGACAAATTAATTCGATTCACCATATCTACTGTGAAT 718

[illegible]

[illegible]

QY 793 TCAGTACCTGTGTCTGGTCCATGATGAGGAAATATGATATATTCACGCTGAT 852
 Db 150 ACTGTTCCTGGTGTCTGGTCCATGATGAGGAAATATGATATATTCACGCTGAT 852
 QY 853 TCATTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 912
 Db 90 GCCTTATTCATGATGAGGAAATATGATATATTCACGCTGAT 912
 QY 913 GTGACAAATATGATGAGGAAATATGATATATTCACGCTGAT 941
 Db 30 CTAATATATGATGAGGAAATATGATATATTCACGCTGAT 941

RESULT 15
 CDS02005 887 bp DNA linear GSS 14-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
 DEFINITION 156L20 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION AL208693.1 GI:7867512
 VERSION AL208693.1
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 887)
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Chausse, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)
 TITLE 2 (bases 1 to 887)
 JOURNAL 2 (bases 1 to 887)
 REFERENCE 2 (bases 1 to 887)
 AUTHORS 2 (bases 1 to 887)

Search completed: October 6, 2002, 07:15:31
 Job time: 912.255 secs

Db 584 TGTTCCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 375 GCCACCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 524 GCCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 435 GAGTGCCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 464 ACCGAGCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 495 AGTGCCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 404 GGCACCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 552 CAAGGAGATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 344 GAAACCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 611 -----AGTATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 284 AACGAGATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 663 ATTCGAGATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 224 GGCACCTTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 720 TATTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 164 TATTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 780 ATTTTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 104 CTTTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 QY 840 TATTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525
 Db 44 CTTTCATTCATGATGAGGAAATATGATATATTCACGCTGAT 525

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 22.0361 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-3_COPY_115_1194

Perfect score: 1080

Sequence: 1 atgacatacaatcccaag.....agaatggtttggaattaa 1080

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_MA:*

1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128.6	11.9	1814	2	US-08-483-151-1
2	128.6	11.9	1814	5	PCT-US96-06427-1
3	128.2	11.9	1086	1	US-07-914-281-12
4	128.2	11.9	1086	1	US-08-393-246-12
5	128.2	11.9	1086	1	US-08-525-058A-12
6	128.2	11.9	1086	4	US-08-696-731-12
7	128.2	11.9	1086	4	US-09-042-531-12
8	128.2	11.9	2042	3	US-09-063-237-2
9	128.2	11.9	2043	1	US-07-914-281-1
10	128.2	11.9	2043	1	US-08-393-246-1
11	128.2	11.9	2043	1	US-08-273-411-4
12	128.2	11.9	2043	1	US-08-525-058A-1
13	128.2	11.9	2043	2	US-08-696-731-1
14	128.2	11.9	2043	4	US-09-042-531-1
15	128.2	11.9	2043	5	PCT-US91-00899-1
16	126.6	11.7	1316	1	US-07-914-281-10
17	126.6	11.7	1316	1	US-08-393-246-10
18	126.6	11.7	1316	1	US-08-525-058A-10
19	126.6	11.7	1316	2	US-08-696-731-10
20	126.6	11.7	1316	4	US-09-042-531-10
21	124.6	11.5	1654	1	US-07-914-281-13
22	124.6	11.5	1654	1	US-08-393-246-13
23	124.6	11.5	1654	1	US-08-525-058A-13
24	124.6	11.5	1654	2	US-08-696-731-13
25	124.6	11.5	1654	4	US-09-042-531-13
26	120	11.1	1068	5	PCT-US91-00899-13
27	120	11.1	1215	5	PCT-US91-00899-5

28	120	11.1	1488	1	US-07-914-281-9	Sequence 9, Appl
29	120	11.1	1488	1	US-08-393-246-9	Sequence 9, Appl
30	120	11.1	1488	1	US-08-525-058A-9	Sequence 9, Appl
31	120	11.1	1488	2	US-08-696-731-9	Sequence 9, Appl
32	120	11.1	1488	4	US-09-042-531-9	Sequence 9, Appl
33	120	11.1	2175	4	US-08-482-073-9	Sequence 9, Appl
34	120	11.1	2861	4	US-08-482-073-10	Sequence 10, Appl
35	120	11.1	3647	1	US-07-914-281-7	Sequence 7, Appl
36	120	11.1	3647	1	US-08-393-246-7	Sequence 7, Appl
37	120	11.1	3647	1	US-08-525-058A-7	Sequence 7, Appl
38	120	11.1	3647	2	US-08-696-731-7	Sequence 7, Appl
39	120	11.1	3647	4	US-09-042-531-7	Sequence 7, Appl
40	120	11.1	3647	5	PCT-US91-00899-4	Sequence 4, Appl
41	116.8	10.8	2134	2	US-08-483-151-3	Sequence 3, Appl
42	116.8	10.8	2134	5	PCT-US96-06427-3	Sequence 3, Appl
43	33.8	3.1	1730	1	US-07-817-920-1	Sequence 1, Appl
44	33.8	3.1	1730	1	US-08-117-006-1	Sequence 1, Appl
45	33.8	3.1	1730	1	US-08-216-594-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-483-151-1
; Sequence 1, Application US/08483151
; Patent No. 5858752
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Holgersson, Jan
; TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,151
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Leach, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/278001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-8906
; TELEFAX: 617/542-5070
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-483-151-1

Query Match 11.9%; Score 128.6; DB 2; length 1814;
Best Local Similarity 52.5%; Pred. No. 1.1e-29;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

QY 269 GCGCCATCTCACAGACCGCTCATTTGTACACAAATCCATGCGGTCCTGATTACACC 328
DB 548 GCGCCGCTGTGATGTCTACACCGAGCGCTGTAGCCAGTGTCTGTGCTGTCTTCCACC 607

Tue Oct 8 10:18:19 2002

us-09-744-748-3_copy_115_1194.rml

Page 2

329 ATAGACATACAGTGGAGTCTGACTAATCTTACCTGAGCAGGCGCAGCCACCTTTGAGA 388
DB ACCGTGAGCTGCAAAACCCGCAATCTCTCACTCCCTGAGCAGAGGCGCAGCAGCAGC 667
389 AATGATTTGATGAAATTTAGTACACCCACTGACACCCCGCAAAAGAGTGACATTAAC 448
DB CTTGGGTCTGGGCTCCATGAGAAATGGCCATTAATACCAATGGCTCCATGCTCCGCG 727
449 ACTGTCAACCTGACTCTAATCTTATGCGCTGATTTAGATATCAAGTGGCTTATGGCT 508
DB GATCTTCAACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
509 TCTGAGGTGAGCACAATTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 568
DB GCTTGGAGCTCTCTCTGTTGGGCTCC--ACATCCCTCCCTACCGGCAAAAGCAGATGGCTG 844
569 GCTGGTGTGAGTAACTGGAACCTGAGCAGTGGGCTGAGTATTAACGAGCTCA 628
DB CCTGGGTGATCAGCAATTTCCAGGAGCGGCGAGCGCTGCAAAAGCTGATACGGGAGCTG 904
629 GCAGAGTATTAATCCACACTATGCGCAAGATTCGAGAAATACGTAACATTAATA 688
DB CCCCTCATCTGAGGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 964
689 ATCTGATTTCCCACTATATCTACTTGAATTTTATCTTCAATTTGAAACCTCAATGCA 748
DB GTCTGCTGCCCTCTTGGCCCGCTTACCGCTTCTACCTGAGCTTGAAGTCAAGCATC 1024
749 AAGATTACATCAGAAAGCTCTA--CAATGCAATTTTGGCTGTTGAGTACCTGTTG 805
DB GGGACTACATCAGTGAAGTCTGCGCAATGCGCTGGGCTGTTGTTGTTGTTGTTGTTGTT 1084
806 TCGGGTCCATCTAGGAAATCTATGAAATTAATCCAGTGTATCATCATCATCATCAT 865
DB CGTGGGACCTCTCTGGGCGCAGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1144
866 TGGAGATTTTAACTCTCCAGTGAAGTAC 896
DB TGGAGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145

RESULT 2
PCT-US96-06427-1
Sequence 1, Application PC/TUS9606427
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06427
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,151
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/278W01
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1814 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US96-06427-1

Query Match 11.9%; Score 128.6; DB 5; Length 1814;
Best Local Similarity 52.5%; Pred. No. 1.1e-29;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

269 GGTGGCATCTCAGACAGACCGCTCATCTGATACAAATCCATGGGCTCTGATTCAC 328
DB GCTGCGCTCTAGTCTACCGGAGCTGCTAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
548 GCTGCGCTCTAGTCTACCGGAGCTGCTAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
DB ACCGTGAGCTCAAAACCCGCAATCTCTCACTCCCTGAGCAGAGGCGCAGCAGCAGCAGCAGC 667
329 ATAGACATCAGTGGAGTCTGACTAATCTTACCTGAGCAGGCGCAGCCACCTTTGAGA 388
DB ACCGTGAGCTCAAAACCCGCAATCTCTCACTCCCTGAGCAGAGGCGCAGCAGCAGCAGCAGC 667
389 AATGATTTGATGAAATTTAGAGTCAACCCACTGACACCCCGCAAAAGAGTGGCATTTGAA 448
DB CTTGGGTCTGGGCTCCATGAGAAATGGCCATTAATACCAATGGCTCCATGCTCCGCG 727
668 GTTGGGTCTGGGCTCCATGAGAAATGGCCATTAATACCAATGGCTCCATGCTCCGCG 727
449 ACTTGTCAACCTGACTTAACTTATGCGCGTATGAGATATCAAGTGGCTTATGGCT 508
DB GATCTTCAACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
728 GATCTTCAACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
DB TCTTACCGGTGAGCACAATCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 844
509 TCTTACCGGTGAGCACAATCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 844
DB TCTTACCGGTGAGCACAATCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 844
788 GTTGGAGCTCTCTCTGTTGGGCTCC--ACATCCCTCACTACCGGCAAAAGCAGATGGCTG 844
569 GCTGGGTGTGAGTAACTGGAACCCCTGAGCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
DB GCTGGGTGTGAGTAACTGGAACCCCTGAGCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
845 CCTGGGTGTGAGTAACTGGAACCCCTGAGCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 904
629 GCAGAGTATTAATCCACACTATGCGCAAGATTCGAGAAATACGTAACATTAATA 688
DB CCCCTCATCTGAGGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 964
905 CCCCTCATCTGAGGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 964
689 ATCTGATTTCCCACTATATCTACTTGAATTTTATCTTCAATTTGAAACCTCAATGCA 748
DB ATCTGATTTCCCACTATATCTACTTGAATTTTATCTTCAATTTGAAACCTCAATGCA 748
965 GTCTGCTGCCCTCTTGGCCCGCTTACCGCTTCTACCTGAGCTTGAAGTCAAGCATC 1024
749 AAGATTACATCAGAAAGCTCTA--CAATGCAATTTTGGCTGTTGAGTACCTGTTG 805
DB TGGAGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
1025 GGGACTACATCAGTGAAGTCTGCGCAATGCGCTGGGCTGTTGTTGTTGTTGTTGTTGTTGTT 1084
806 TCGGGTCCATCTAGGAAATCTATGAAATTAATCCAGTGTATCATCATCATCATCATCATCATCATCAT 865
DB CGTGGGACCTCTCTGGGCGCAGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1144
1085 CGTGGGACCTCTCTGGGCGCAGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1144
866 TGGAGATTTTAACTCTCCAGTGAAGTAC 896
DB TGGAGCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145

RESULT 3
US-07-914-281-12
Sequence 12, Application US/07914281
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/914, 281
 FILING DATE: 19920720
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-07-914-281-12

Query Match 11.9%; Score 128.2; DB 1; Length 1086;
 Best Local Similarity 53.0%; Pred. No. 1.2e-29;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

271 TGGCATCTCAACAGAGCGCTATTGTACACAAATCCCATCGCTGATTCACCAT 330
 271 TGGCATCTCAACAGAGCGCTATTGTACACAAATCCCATCGCTGATTCACCAT 330
 271 TGGCATCTCAACAGAGCGCTATTGTACACAAATCCCATCGCTGATTCACCAT 330
 331 AGAGACATCAGCTGGGA--TCTGACTAACTTACCTCAGAGGCCAGCCCTTTGAG 387
 331 TGGGAATATCTGTCCAAACCTTAAGTACAGCGCTCCCACTTCCCGAGGCGGAG 390
 388 AATGATTTGATGATTAAGTACAGCGCTCCCACTTCCCGAGGCGGAG 447
 391 CGCTGATCTGTTCAACTTGGAGCACCCTTAACCTGCGAGCACTGGAAGCCTGGAC 450
 448 CACTTGTCAACCTGACTTAACTTACCTGATTCAGATTCAGATTCAGATTCAG 507
 451 AGATACCTTCAATCTCAACATGCTCTACCGGAGGAGCTCCGACATCTTACGCG 510
 508 TTTTACAGGTGAGCAAAATCCCTTTG-----TGTTTAAGTGGCCAGCAGGAG 558
 511 TGGCTGAGCGCGGTGTCGCGGCGAGCTGCCACCAACCGCTCAACTCTCTGCGCAG 570
 559 AAGTGTGCTGCTGGTGTGATTAAGTACAGCGCTGAGCAGCAGGCGTCAAGTATTA 618
 571 GACTGTGCTGCTGGGCGGTGTCAACTGAGAACCGGACTCAGCGAGGCGCTCTACTAC 630
 619 AACGAGCTCAGCAAGATTTGAATTCACACCTATGCGCAGCAGCTTGGAGATACGTG 678
 631 CAGAGCTGAGGCTCATCTCAAGGTGAGAGGTACGAGAGCTCCACACA--AGCCCTG 687
 679 AACGATTAATAATCTGATTCCTCCACCATATCTACTTTGAATTTTATCTTTGAAAC 738
 688 CCAAGGAGGAGCATGATGAGAGCTGTCCCGGTACAAAGTTCTACCTGGGCTTCGAGAC 747
 739 TCAATTCACAAAGATTAATCAGAGAAAGCTCT--ACAATGATTTTGGCTGGTTCA 795
 748 TCTTGTGACCCCGCATACATCAGAGAGAGAGCTGTGAGAGAGCCCTTGAGGCTG 807

796 GTACCTTTCTCTGCTGCTCATAGGAAACTATGATTAATTCAGCTGATTC 855
 808 GTGCCCTGTGGTGGGCGCCAGACAGCAAGTACAGAGGTTCCTCCACCCGAGCC 867
 856 TTCTATATGTGGAAGATTTTAATCTCCAGAGAGTACCAAAATCTGGAAGAGTT 915
 868 TTATCCAGCTGGAGCACTTCAGAGAGCCCAAGAGCTGCGGCTACTGACAGAGCTG 927

916 GACAAAACATTAAGTGTACCTTACTTACTTACTGAGGAGAA 960
 928 GACAGAGCAGCAGCCCGCTACTGAGCTACTTGGCTGGCGGAG 972

RESULT 4
 US-08-393-246-12
 ; Sequence 12; Application US/08393246
 ; Patent No. 5595900

GENERAL INFORMATION:
 APPLICANT: LOWE, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 TITLE OF INVENTION: GLYCOPOLYMER, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:

CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994
 APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-393-246-12

Query Match 11.9%; Score 128.2; DB 1; Length 1086;
 Best Local Similarity 53.0%; Pred. No. 1.2e-29;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

271 TGGCATCTCAACAGAGCGCTATTGTACACAAATCCCATCGCTGATTCACCAT 330
 271 TGGCATCTCAACAGAGCGCTATTGTACACAAATCCCATCGCTGATTCACCAT 330
 271 TGGCATCTCAACAGAGCGCTATTGTACACAAATCCCATCGCTGATTCACCAT 330
 331 AGAGACATCAGCTGGGA--TCTGACTAACTTACCTCAGAGGCCAGCCCTTTGAG 387

Tue Oct 8 10:18:19 2002

US-09-744-748-3_COPY_115_1194.rml

Db 331 TGGGATATCATGTCACACCCCTAAGTCACGCCCTCCACACCTTCCCGAGGCCGAGGCGAG 390
OY 388 AAGTGGATTTGATGATTAATTTAGAGTACACCCGACACACCCGAGGCGAGGCGAGGCGAG 390
Db 391 CGGTGGATGCTGTTCACTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 447
OY 448 CAGTGGTTCACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 450
Db 451 AGATACCTTCAATCTCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 450
OY 508 TTTTGAAGGCTGACGACCAATGCTTGG-----TGTTAAGTGGCAAGCAAGAGAG 558
Db 511 TGGCTGGAGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
OY 559 AAGTGG 558
Db 571 GAGCTGG 570
OY 619 AAGGAGCTCAGACCAAGATTTGAATTCACACCTGACGACGACGACGACGACGACGACGACGAC 630
Db 631 CAGACCTGCGAGGCTCATCTCAAGGTGACGAGTGAAGGCTGCGACGACGACGACGACGACGAC 678
OY 679 AAGCATTAATTTGATGATTAATTTAGAGTACACCCGACACCCGAGGCGAGGCGAGGCGAG 687
Db 688 CCGAAGGGGACCATGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 738
OY 739 TCAATTCACCAAGATTTAGATTAATTTAGAGTACACCCGACACCCGAGGCGAGGCGAGGCGAG 747
Db 748 TCTTTCACACCCGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 795
OY 796 GTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Db 808 GTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
OY 856 TTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
Db 868 TTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
OY 916 GACAAAACAAATTAAGTTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
Db 928 GACAAAGACCAAGCCGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972

RESULT 5
US-08-525-058A-12
Sequence 12, Application US/08525058A
Patent No. 5,720,420
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
NUMBER OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: U.S.A.
22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/525,058A
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalley, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55
TELEPHONE: (703) 521-4500
TELEFAX: (703) 521-4500
INFORMATION FOR SRO: OPAT UR
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
US-08-525-058A-12
Query Match
Best Local Similarity 11.9%; Score 128.2; DB 1; Length 1086;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4.

Db 271 TGGCATTCACCAAGACCGCTGATTTGACAAACAATGCCATGCGGTCTGATGACCAT 330
OY 271 TGGCATTCACCGACCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 330
Db 331 AAGACATCAGTGGGA---TGTGACTTACTTACCTGACGACGACGACGACGACGACGACGACGAC 387
OY 388 AAGTGGATTTGATGATTAATTTAGAGTACACCCGACACCCGAGGCGAGGCGAGGCGAG 390
Db 391 CGGTGGATGCTGTTCACTGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 447
OY 448 CAGTGGTTCACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 450
Db 451 AGATACCTTCAATCTCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 507
OY 508 TTTTGAAGGCTGACGACCAATGCTTGG-----TGTTAAGTGGCAAGCAAGAGAG 558
Db 511 TGGCTGGAGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
OY 559 AAGTGG 570
Db 571 GAGCTGG 618
OY 619 AAGGAGCTCAGACCAAGATTTGAATTCACACCTGACGACGACGACGACGACGACGACGACGAC 630
Db 631 CAGACCTGCGAGGCTCATCTCAAGGTGACGAGTGAAGGCTGCGACGACGACGACGACGACGAC 678
OY 679 AAGCATTAATTTGATGATTAATTTAGAGTACACCCGACACCCGAGGCGAGGCGAGGCGAG 687
Db 688 CCGAAGGGGACCATGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 738
OY 739 TCAATTCACCAAGATTTAGATTAATTTAGAGTACACCCGACACCCGAGGCGAGGCGAGGCGAG 747
Db 748 TCTTTCACACCCGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 795
OY 796 GTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Db 808 GTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
OY 856 TTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
Db 868 TTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
OY 916 GACAAAACAAATTAAGTTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
Db 928 GACAAAGACCAAGCCGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972

RESULT 6
US-08-696-731-12
Sequence 12, Application US/08696731
Patent No. 5,955,347
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.

Tue Oct 8 10:18:19 2002

us-09-744-748-3_copy_115_1194.rn1

STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-12

Query Match
Best Local Similarity 11.9%; Score 128.2; DB 4; Length 1086;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 271 TGGCATTCACAAAGACCGCTCATTTGTAACAAATTCATGCGTCTGATTCACCAT 330
DB 271 TGGCATTCACAAAGACCGCTCATTTGTAACAAATTCATGCGTCTGATTCACCAT 330
QY 331 AGAGCATCAGCTGGGA---TGTACTTACTTACTTACTTACTTACTTACTTACTTACTT 330
DB 331 AGAGCATCAGCTGGGA---TGTACTTACTTACTTACTTACTTACTTACTTACTTACTT 330
QY 331 TGGGATATCATGTCACACCTTAAAGTACAGCCCTCCACCTCCGAGGCGGAG 390
DB 331 TGGGATATCATGTCACACCTTAAAGTACAGCCCTCCACCTCCGAGGCGGAG 390
QY 388 AATATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
DB 388 AATATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 391 CGCTGATCTGTTCACTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 450
DB 391 CGCTGATCTGTTCACTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 448 CACTTTCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
DB 448 CACTTTCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 451 AGATCTTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 507
DB 451 AGATCTTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 507
QY 508 TCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
DB 508 TCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
QY 511 TGGCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
DB 511 TGGCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
QY 559 AAGTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 559 AAGTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
QY 571 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
DB 571 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
QY 619 AAGGACTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
DB 619 AAGGACTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
QY 631 CAGAGCTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
DB 631 CAGAGCTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
QY 679 AAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
DB 679 AAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
QY 688 CCCAAGGGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 688 CCCAAGGGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
QY 739 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
DB 739 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
QY 748 TCCCTGACCCGACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 807
DB 748 TCCCTGACCCGACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 807
QY 796 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
DB 796 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
QY 808 GTCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
DB 808 GTCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
QY 856 TTTCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
DB 856 TTTCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
QY 916 GACAAAACATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
DB 916 GACAAAACATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
QY 928 GACAAAGACACCGCGCTACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
DB 928 GACAAAGACACCGCGCTACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972

RESULT 8
US-09-063-237-2
Sequence 2, Application US/09063237
Patent No. 612426
GENERAL INFORMATION:
APPLICANT: McPher, Roger P.
APPLICANT: Cummings, Richard D.
TITLE OF INVENTION: O-Glycan Inhibitors of Selectin Mediated
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta

STATE: Georgia
COUNTRY: US
ZIP: 30305
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/063,237
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
PUBLICATION NUMBER: 08/649,802
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
TELEPHONE/DOCKET NUMBER: OMRI110CIP7
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8794
INFORMATION FOR SDO NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-063-237-2

Query Match
Best Local Similarity 11.9%; Score 128.2; DB 3; Length 2042;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
QY 271 TGGCATTCACAAAGACCGCTCATTTGTAACAAATTCATGCGTCTGATTCACCAT 330
DB 271 TGGCATTCACAAAGACCGCTCATTTGTAACAAATTCATGCGTCTGATTCACCAT 330
QY 331 AGAGCATCAGCTGGGA---TGTACTTACTTACTTACTTACTTACTTACTTACTTACTT 330
DB 331 AGAGCATCAGCTGGGA---TGTACTTACTTACTTACTTACTTACTTACTTACTTACTT 330
QY 331 TGGGATATCATGTCACACCTTAAAGTACAGCCCTCCACCTCCGAGGCGGAG 390
DB 331 TGGGATATCATGTCACACCTTAAAGTACAGCCCTCCACCTCCGAGGCGGAG 390
QY 388 AATATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
DB 388 AATATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 391 CGCTGATCTGTTCACTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 450
DB 391 CGCTGATCTGTTCACTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 448 CACTTTCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
DB 448 CACTTTCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 451 AGATCTTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 507
DB 451 AGATCTTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 507
QY 508 TCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
DB 508 TCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
QY 511 TGGCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
DB 511 TGGCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
QY 559 AAGTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 559 AAGTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
QY 571 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
DB 571 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
QY 619 AAGGACTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
DB 619 AAGGACTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
QY 631 CAGAGCTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
DB 631 CAGAGCTCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
QY 679 AAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
DB 679 AAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
QY 688 CCCAAGGGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 688 CCCAAGGGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
QY 739 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
DB 739 TCAATTCACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
QY 748 TCCCTGACCCGACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 807
DB 748 TCCCTGACCCGACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 807
QY 796 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
DB 796 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855

Db 880 GTGCCCCGTGCTGCTGGCCCCAGACAGCACTACGAGAGGTTCTCCACCCGACGCC 939
Qy 856 TTGATTCATGTGGAGATTTTAACTCTCCAGTAGTACAAATATCTGAAGAAGTT 915
Db 940 TTGATTCATGTGGAGATTTTAACTCTCCAGTAGTACAAATATCTGAAGAAGTT 915
Qy 916 GACAAAACATATGATTGTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 960
Db 1000 GACAAAGACCACGCCGCTACTGACTTACTTACTTACTTACTTACTTACTTACTT 1044

RESULT 9

US-07-914-281-1
; Sequence 1, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,281
; FILING DATE: 19920720
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2043 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ANTI-SENSE: NO
; US-07-914-281-1

Query Match 11.9%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1.6e-29;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

Qy 271 TGCATCTCACAAGACCGCTCATTTGACAAATCCATGCGTCTGATTCACCAT 330
Db 343 TGCACATCATGCGACGCGCAAGGTGACCAAGCAGACAGGTCATGTCACAC 402
Qy 331 AGAGCATAGCTGGGA---TCTGACTACTTACTTACCTCAGAGCCGCGCATTTGAG 387
Db 403 TGGGATATCATGTCACACCTTAAGTACGCTCCGCTCCCGGAGGCGGAGG 462
Qy 388 AATGATTTGATGATTAATTAAGTACACCCACTACACCCCAAAAGAGTGGCATTTGA 447
Db 463 CGCTGATCTGTTCAACTTGGAGCCACCCCTTAAGTACGACGACCTGGAAGCCCTGGAG 522

Qy 448 CACTGTTCACCTGACTTAACTATGCGCGTGTATTCAGATATCCAGTGCCTATGAGC 507
Db 523 AGATTTCTGATTCATGTCACATGTCACGAGGACTCCGACATCTTCACGCCCTAGGC 582
Qy 508 TTCTTACGCTGAGCACAATCCCTTTG-----TCTTGAATGCCAAGAGAG 558
Db 583 TGGCTGAGCCCGGTGTCGCGCCAGCCGCTCCACCCGCTCAACCTCTCGGCGAAGACC 642
Qy 559 AGTTGTGTGCTGGTGTGATGATTAAGTACAGACCCCTGAGCATGCGAGGTCAGATATAC 618
Db 643 GAGCTGTGCTGCGCGGCTGTCACACTGGAAGCCGAGCTACGCCAGGCTGCTACTAC 702
Qy 619 AACGACTCAGCAAGAGATTTGAATTCACACCTATGCGCAAGCATTCGAGAAATACGTG 678
Db 703 CAGAGCTCAGAGCTCATCTCAAGGTGAGCTGTACGACGCTCCACA---AGCCCCG 759
Qy 679 AACGATTAATATCTGATTTCCACATATCTACTTGTAAATTTTATCTTATTCATTTGAAAC 738
Db 760 CCCAAGGGAGCATATGAGAGAGCGTCCGCTACCAAGTTCTACTGCTTCCGAGAAC 819
Qy 739 TCATTCACAAGATATACATACAGAAAGCTCT---ACATGCAATTTTGGCTGTTCA 795
Db 820 TCTTGTACACCCGACTACATACACGGAAGCTGTGAGAGAGCCCTGAGGCTGAGCC 879
Qy 796 GTACTGTGTCCTGGCTCATCTAGGAAACTATGAGAATTAATATTCACAGCTGATTC 855
Db 880 GTGCCCCGTGCTGCTGGCCCCAGACAGCACTACGAGGTTCTGCGCCAGCCGCGC 939
Qy 856 TTGATTCATGTGGAGATTTTAACTCTCCAGTAGTACAAATATCTGAAGAAGTT 915
Db 940 TTGATTCATGTGGAGATTTTAACTCTCCAGTAGTACAAATATCTGAAGAAGTT 915
Qy 916 GACAAAACATATGATTGTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 960
Db 1000 GACAAAGACCACGCCGCTACTGACTTACTTACTTACTTACTTACTTACTTACTT 1044

RESULT 10

US-08-393-246-1
; Sequence 1, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451

Tue Oct 8 10:18:19 2002

us-09-744-748-3_copy_115_1194.rni

Page 8

REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ANTI-SENSE: NO
US-08-393-246-1

Query Match 11.9%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1.6e-29;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

OY 271 TGCATCTCAACAAGACGCGCTCATTTGTACAAATAATCCATGCGGCTCTGATTCACCAT 330
DB 343 TGCACATCACTGCGCGAGCGAGGTACCCACAGCGAGACAGCGTCTATGTCACACAC 402
OY 331 AGACATCACTGCGCGAGCGAGGTACCCACAGCGAGACAGCGTCTATGTCACACAC 387
DB 403 TGGATATTCATGTCACACCTTAAGTACGCGCTCCACCTTCCCGAGGCGCGACAGGGCAG 462
OY 388 AATGATTTGGATGAATTTAGAGTACCCACTCAGCCCGCCCAAGAGAGTGGCATTTGAA 447
DB 463 CGCTGATCTGATCACTGAGAGCCACCCCTTAATGTCAGACCTGGAAGCCCTTGAC 522
OY 448 CACTGTTCACCTGACTCTAATTTATGCGCGTGTATGATGATTCAGATTCAGATTCAGC 507
DB 523 AGATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 582
OY 508 TTCTTGACGCTGAGCAACAATCCCTTTG-----TGTGGAAGTCCCAAGCAAGAG 558
DB 583 TGGCTGAGCGCGGTGTCGCGCGAGCGCTGCCACCGCATCCTCGGCGCAAGAGAC 642
OY 559 AAGTGTGCTGCTGGGTTGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 618
DB 643 GAGTGTGCTGCTGGGTTGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 702
OY 619 AAGCAGCTCAGCAAGAGTATGAAATCCACACCTATGCGCAAGATTCGAGAAATAGTG 678
DB 703 CAGAGCTCAGCAAGAGTATGAAATCCACACCTATGCGCAAGATTCGAGAAATAGTG 759
OY 679 AAGCAGCTCAGCAAGAGTATGAAATCCACACCTATGCGCAAGATTCGAGAAATAGTG 738
DB 760 CCGAAGGAGGACATGATGAGAGAGCGTCCCGGTACAAAGTCTTACCTGCGCTTGAGAGAC 819
OY 739 TCAATTCACAAAGATTTACATCAGCAAGAGAGCTCT--ACAATGCAATTTTGGCTGTGCA 795
DB 820 TCTCTGACCGCGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 879
OY 796 GTACCTGTGCTGCGGTTGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 855
DB 880 GTGCGCGGTGCTGCGGTTGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 939
OY 856 TTTCAATTCAGCAAGAGTATGAAATCCACACCTATGCGCAAGATTCGAGAAATAGTG 915
DB 940 TTTCAATTCAGCAAGAGTATGAAATCCACACCTATGCGCAAGATTCGAGAAATAGTG 999
OY 916 GACAAAAACAATAGTTGATCTTACTTAACTGAGAGAG 960
DB 1000 GACAAAGAGCAAGCGCGCTCACTTCTGCTGCGGAG 1044

RESULT 11
US-08-273-411-4
Sequence 4, Application us/08273411
Patent No. 5625124
GENERAL INFORMATION:

APPLICANT: Falk, Per
APPLICANT: Gordon, Jeffrey I.
TITLE OF INVENTION: Animal Model for Gastro-Intestinal
DISEASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4330
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,411
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: R7106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 815-6508
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc. feature
LOCATION: 1..361
OTHER INFORMATION: /note="Nucleotides 73 through 1158 encode the GDP-L-fuco
PUBLICATION INFORMATION:
AUTHORS: Kulkowska-Latallo, et al.
JOURNAL: Genes & Development
VOLUME: 4
PAGES: 1288-1303
DATE: 1990
RELEVANT RESIDUES IN SEQ ID NO: 4: FROM 1 TO 2043

Query Match 11.9%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1.6e-29;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

OY 271 TGCATCTCAACAAGACGCGCTCATTTGTACAAATAATCCATGCGGCTCTGATTCACCAT 330
DB 343 TGCACATCACTGCGCGAGCGAGGTACCCACAGCGAGACAGCGTCTATGTCACACAC 402
OY 331 AGACATCACTGCGCGAGCGAGGTACCCACAGCGAGACAGCGTCTATGTCACACAC 387
DB 403 TGGATATTCATGTCACACCTTAAGTACGCGCTCCACCTTCCCGAGGCGCGACAGGGCAG 462
OY 388 AATGATTTGGATGAATTTAGAGTACCCACTCAGCCCGCCCAAGAGAGTGGCATTTGAA 447
DB 463 CGCTGATCTGATCACTGAGAGCCACCCCTTAATGTCAGACCTGGAAGCCCTTGAC 522
OY 448 CACTGTTCACCTGACTCTAATTTATGCGCGTGTATGATGATTCAGATTCAGATTCAGC 507
DB 523 AGATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 582
OY 508 TTCTTGACGCTGAGCAACAATCCCTTTG-----TGTGGAAGTCCCAAGCAAGAG 558
DB 583 TGGCTGAGCGCGGTGTCGCGCGAGCGCTGCCACCGCATCCTTCTGCGCAAGAGAC 642


```

; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,731
; FILING DATE: 14-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/393,246
; FILING DATE:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SRO ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2043 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ANTI-SENSE: NO
; US-08-696-731-1

Query Match 11.9%; Score 128.2; DB 2; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1.6e-29;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

OY 271 TGGCATTCACACAGACGCTGATGTACAAACAATCCCATGGGCTGATTCACAT 330
DB 343 TGGCACAATCACTGCCGACCCGAGGTGTACCCACAGCAGACACGCTGACACAC 402
OY 331 AGACACATGAGCTGGGA---TCGACTACTACTTACCTCAGCAGCCACCCCTTTAG 387
DB 403 TGGGATATCATGTCCAAACCTTAAGTACAGCTCCACCTTCCCGAGGCGAGGGGAG 462
OY 388 AATGATTTGGATGAATTTAGAGTACACCCACTCACACCCCAAAAGAGTGGCATGAA 447
DB 463 CGCTGGATCTGTTCAACTGAGACCCACCCCTTAAGTGTGACACCTGGAACCCCTGAC 522
OY 448 CACTTGTTCACACTGACTACTTACTATGCGCGTATTCAGATATCCAGTGCCTATGAC 507
DB 523 AGATACCTTCATCTCACCATGTCTTACCGCAGGACTCCGACATCTTCACGCTTACGCG 582
OY 508 TTCTTGAAGGAGACACAATATCCCTTTG-----TGTGAGAGTCCCAAGAGAG 558
DB 583 TGGCTGGAGCGGTGTGTGCTGGGCGACGCTGCCACCCACCCCTCAACCTCTCGGCGAAGAC 642
OY 559 AAGTTGCTGTCTGGTGTGAGTAACTGGAACCTGACATGCCAGGCTCAAGTATTAC 618
DB 643 GAGCTGGGCTGTGGGCTGTCCAACTGGAACCCGAGTCAAGCGTGGCTACTAC 702
OY 619 AAGAGCTCAGCAAGATATTTGAATTCACACCTATGGCAAGCATTTGGAATATACGTG 678
DB 703 CAGAGCCTGCAGGCTCATCTCAAGGTGAGCTGTACGAGCGTCCACACA---AGCCCTG 759
OY 679 AACGATATAATCTGATCCACCATATCTACTTGAATTTATCTTCAATTGAATAC 738
DB 760 CCNAAGGAGACATGATGAGAGCGCTGTCCCGGTACAAATTTACTGCGCTTGAGAAC 819
OY 739 TCATTCACAAAGATTCATCATACAGAAAGCTCT---ACAATGCAATTTTGGCTGTGCA 795

```

```

DB 820 TCCCTGACCCCGACTATCATCCAGAGAGCTGTGAGAGAACCCCTGGAGGCTGGGCC 879
OY 796 GTACCTGTTGCTCCGATCTAGGAGAAACATATGAAATATATTCAGCTATCA 855
DB 880 GTGCCGTGTGTGTGGGCGCCAGCAGACACTAGAGAGTTCCTGCGACCCGACGCC 939
OY 856 TGCATTCATGCGAAGATTTTAATCTTCCAGAGAGTGTAGCAAAATATCGAAGAGTT 915
DB 940 TATATCAGGTGAGACTTCCAGACCCCAAGAGCTGACCGGTGCTGCGAGAGCTG 999
OY 916 GACAAAACATATAGTTACTTACTTACTTACTGAGAAAG 960
DB 1000 GACAAGAGCACGCGGCTGACTGACTGCTGCTGCGGAG 1044

RESULT 14
US-09-042-531-1
; Sequence 1, Application US/09042531
; Patent No. 6268193
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,531
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SRO ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2043 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ANTI-SENSE: NO
; US-09-042-531-1

Query Match 11.9%; Score 128.2; DB 4; Length 2043;
Best Local Similarity 53.0%; Pred. No. 1.6e-29;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

OY 271 TGGCATTCACACAGACGCTGATGTACAAACAATCCCATGGGCTGATTCACAT 330

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:39:45 ; Search time 1146.59 Seconds
(without alignments)
16699.717 Million cell updates/sec

Title: US-09-744-748-3_COPY_280_1194

Perfect score: 915
Sequence: 1 acaaaaactgattatttaa.....agaatggttcgaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	915	100.0	2139	10	AB015426 Mus muscu
2	856	93.6	1128	10	AF345993 Rattus no
3	852.8	93.2	2156	10	AB049819 Rattus no
4	825.4	90.2	1705	10	AF230460 Cricetus
5	796.6	87.1	2501	9	AF238701 Homo sapi
6	793.4	86.7	80247	9	AL512406 Human DNA
7	791.8	86.5	3019	9	AB023021 Homo sapi
8	668.4	73.0	71396	2	AC100674 Mus muscu
9	589.2	64.4	1080	5	AB035905 Gallus ga
10	589.2	29.3	1567	5	AB023627 Danio rer
11	268.2	24.9	1832	5	AB023628 Danio rer
12	228.2	24.1	1264	5	AF288369 Gallus ga
13	220.6	22.6	1092	5	AB035907 Xenopus l
14	206.6	19.1	2657	5	GGU73678 Gallus gall
15	175.2	17.2	1429	10	CGU78737 Cricetus
16	157	17.2	1462	10	AF090450 Cricetus
17	157	17.2	1462	10	AF090450 Cricetus
18	155.4	17.0	1718	10	AF090449 Cricetus
19	145.6	15.9	71396	2	AC100674 Mus muscu
20	135.6	14.8	1044	5	AB035908 Gallus ga
21	134.8	14.7	2293	4	BTAI34PTG Bos tauru
22	134.8	14.7	12607	4	BTAI32772 Bos tauru
23	134.2	14.7	1158	10	AB039126 Mus muscu
24	133.4	14.6	1113	9	AF345884 Salimeli s
25	133	14.5	1119	9	AF345881 Macaca mu
26	132.6	14.5	1170	10	AB039127 Mus muscu
27	132.6	14.5	1170	10	AB039131 Mus muscu
28	132.6	14.5	1615	10	MMU33457 Mus muscu
29	131.4	14.4	1170	10	AB039133 Mus spici
30	131	14.3	1170	10	AB039124 Mus muscu
31	131	14.3	1170	10	AB039125 Mus muscu
32	131	14.3	1170	10	AB039128 Mus muscu
33	131	14.3	1170	10	AB039129 Mus muscu
34	131	14.3	1170	10	AB039130 Mus muscu
35	131	14.3	1170	10	AB039132 Mus muscu
36	130.8	14.3	1302	10	AB049938 Rattus no
37	130.8	14.3	16946	2	AC097778 Rattus no
38	130	14.2	1613	4	BTAI32775 Bos tauru
39	130	14.2	1619	4	BTAI32776 Bos tauru
40	130	14.2	1644	4	BTAI32773 Bos tauru
41	129.8	14.2	1873	4	BTAI32774 Bos tauru
42	129.8	14.2	1162	9	PTI14033 Bos tauru
43	129.8	14.2	80893	9	AC024592 Homo sapi
44	129.8	14.2	187488	2	AC010454 Homo sapi
45	128.8	14.1	169893	2	AC021165 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE
AB015426	AB015426	Mus musculus Put9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.	AB015426	AB015426.1	GI:3702718	Fut9; alpha1,3-fucosyltransferase IX.	Mus musculus	1 (sites)	Kudo,T., Ikehara,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K. and Narimatsu,H.	Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mFuc-PIX, that synthesizes the Lewis x (CD15) epitope in brain and kidney	J. Biol. Chem. 273 (41), 26729-26738 (1998)	98434588	2 (bases 1 to 2139)	Kudo,T. and Narimatsu,H.	Direct Submision
AB015426	AB015426	Mus musculus Put9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.	AB015426	AB015426.1	GI:3702718	Fut9; alpha1,3-fucosyltransferase IX.	Mus musculus	1 (sites)	Kudo,T., Ikehara,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K. and Narimatsu,H.	Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mFuc-PIX, that synthesizes the Lewis x (CD15) epitope in brain and kidney	J. Biol. Chem. 273 (41), 26729-26738 (1998)	98434588	2 (bases 1 to 2139)	Kudo,T. and Narimatsu,H.	Direct Submision

JOURNAL Submitted (09-JUN-1998) Takaaki Kudo, Institute of Life Science,
Soka University, Division of Cell Biology, 1-236 Tangi-cho,
Hachioji, Tokyo 192-8577, Japan (E-mail: tkudo@soka.ac.jp,
Tel:+81-426-91-2495, Fax:+81-426-91-9315)

FEATURES
source
1. .2139
/organism="Mus musculus"
/db_xref="taxon:10090"
115. .1194
/gene="Fut9"
115. .1194
/gene="Fut9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="PAA33522.1"
/db_xref="GI:3702719"
/translation="MSTSKGIIIRPLIVCIILGCFMACLLIYIKPTNSWVSPMESA
SSVLKMKFEFSTKDYFNETTILVWVPGQFDLTSCQAFNIOGCHLTDRSLYRK
SHAVLIHHRDISMDLTNLPQAPRPQKIMWNLSPTPTPOKSGIHLPLTYRR
DSDIQVPGFLVTSFVEVPEVPSKEKLYCWMNNEBARVYVNELSKSEIHTY
GQAFGEVYNDKNIPISTCKEFLSPENSIRHDIYTERLYNAFLAGSVPVLPSPREN
YENYIPADSFTHVEDFNSPELAKYKLEVDKNNKLYLSYFWMRKDFVNLPRWESHA
CLADHYKRHOEYKSVGNLEKMFVN"
2139

polya_site
/note="31 a nucleotides"
BASE COUNT 684 a 395 c 368 g 692 t
ORIGIN

Query Match 100.0%; Score 915; DB 10; Length 2139;
Best Local Similarity 100.0%; Pred. No. 9.7e-233;
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ACAAACGATATTTTAAAGAACTACATCTCGTGTGGTATGCGCATTTGGGAG 60
Db 280 ACAAACGATATTTTAAAGAACTACATCTCGTGTGGTATGCGCATTTGGGAG 339
Oy 61 ACCTTGACCTTACATCTCGCAGCAAGCATGTTCAATATCCAGAGGTGCATCTCACACA 120
Db 340 ACCTTGACCTTACATCTCGCAGCAAGCATGTTCAATATCCAGAGGTGCATCTCACACA 399
Oy 121 GACCGCTCATTTGACAAACAATCCAGCGGTCTCGATTCACCATAGACATCAGCTGG 180
Db 400 GACCGCTCATTTGACAAACAATCCAGCGGTCTCGATTCACCATAGACATCAGCTGG 459
Oy 181 GATCGACTAACTTACCTAGAGAGCCGACCCCTTTGAGAAATGATTTGATGAAT 240
Db 460 GATCGACTAACTTACCTAGAGAGCCGACCCCTTTGAGAAATGATTTGATGAAT 519
Oy 241 TTAGAGTACCCACACTACACCCGCCAAGAAGTGGCATTCACCTTGTTCACCTGACT 300
Db 520 TTAGAGTACCCACACTACACCCGCCAAGAAGTGGCATTCACCTTGTTCACCTGACT 579
Oy 301 CTAACTTATGCGCGTATTCAGATATCCAGTGCCTTATGCTTGTGACGGTGAAGACA 360
Db 580 CTAACTTATGCGCGTATTCAGATATCCAGTGCCTTATGCTTGTGACGGTGAAGACA 639
Oy 361 AATGCCCTTGTGTTGAATGCGCAGCAAGAGTGGTGGCTGGCTGTGAGTAAAC 420
Db 640 AATGCCCTTGTGTTGAATGCGCAGCAAGAGTGGTGGCTGTGAGTAAAC 699
Oy 421 TGGAACTGAGCATGCCAGGGTCAAGTATTACAGAGCTCAGCAAGAGTATTGAAATC 480
Db 700 TGGAACTGAGCATGCCAGGGTCAAGTATTACAGAGCTCAGCAAGAGTATTGAAATC 759
Oy 481 CACACCTATGCGCAAGCATTCGAGATACGATATAAATATGATTTCCACCATTA 540
Db 760 CACACCTATGCGCAAGCATTCGAGATACGATATAAATATGATTTCCACCATTA 819
Oy 541 TCGACTTGTAAATTTATCTTCAATTTGAATAATCATATGACAAATATACATCAACA 600
Db 820 TCGACTTGTAAATTTATCTTCAATTTGAATAATCATATGACAAATATACATCAACA 879
Oy 601 AAGCTTACATGCAATTTTGGCTGCTCAGTACCTGCTGGGCTCATCTGAGGAA 660

Db 880 AAGCTTACATGCAATTTTGGCTGCTCAGTACCTGCTGGGCTCATCTGAGGAA 939
Oy 661 AACTATGAGAAATTAATTCACAGCTGATTCATTCATGATGGAAGATTTTAACTCTCC 720
Db 940 AACTATGAGAAATTAATTCACAGCTGATTCATTCATGATGGAAGATTTTAACTCTCC 999
Oy 721 AGTGAAGTGAAGAAATATCTGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 1000 AGTGAAGTGAAGAAATTAATTCACAGCTGATTCATTCATGATGGAAGATTTTAACTCTCC 1059
Oy 781 TTTAATCTGAGAGAAATTTTAACTGATTAACCTGATTCACAGCTGATTCATTCATGAT 840
Db 1060 TTTAATCTGAGAGAAATTTTAACTGATTAACCTGATTCACAGCTGATTCATTCATGAT 1119
Oy 841 CTGGCATGCCAGCATTAATAAGGATCAAGATATTAAGTCTGTGTGATTAATTAAGAGAA 900
Db 1120 CTGGCATGCCAGCATTAATAAGGATCAAGATATTAAGTCTGTGTGATTAATTAAGAGAA 1179
Oy 901 TGGTTTGAAGATTA 915
Db 1180 TGGTTTGAAGATTA 1194

RESULT 2
AF345993 1128 bp mRNA linear ROD 07-MAR-2001
LOCUS AF345993
DEFINITION Rattus norvegicus alpha1,3-fucosyltransferase IX (Fut9) mRNA,
complete cds.
ACCESSION AF345993
VERSION AF345993.1 GI:13242183
KEYWORDS
SOURCE Norway rat.
ORIGIN Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1128)
Baboval,T., Henton,T., Kinnally,E. and Smith,F.I.
Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX)
and comparison of the expression of fuc-TIY and fuc-TIX genes
during rat postnatal cerebellum development
J. Neurosci. Res. 62 (2), 206-215 (2000)
1120213
JOURNAL PUBMED
REFERENCE 2 (bases 1 to 1128)
Smith,F.I. and Baboval,T.
Direct Submission
Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver
Center, 200 Trapelo Rd., Waltham, MA, USA
JOURNAL LOCATION/Qualifiers
1. .1128
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1. .1128
/gene="Fut9"
49. .1128
/gene="Fut9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="AAK16591.1"
/db_xref="GI:13242184"
/translation="MSTSKGIIIRPLIVCIILGCFMACLLIYIKPTNSWVSPMESA
SSVLKMKFEFSTKDYFNETTILVWVPGQFDLTSCQAFNIOGCHLTDRSLYRK
SHAVLIHHRDISMDLTNLPQAPRPQKIMWNLSPTPTPOKSGIHLPLTYRR
DSDIQVPGFLVTSFVEVPEVPSKEKLYCWMNNEBARVYVNELSKSEIHTY
GQAFGEVYNDKNIPISTCKEFLSPENSIRHDIYTERLYNAFLAGSVPVLPSPREN
YENYIPADSFTHVEDFNSPELAKYKLEVDKNNKLYLSYFWMRKDFVNLPRWESHA
CLADHYKRHOEYKSVGNLEKMFVN"

BASE COUNT 93 a 256 c 228 g 320 t
ORIGIN

Query Match 93.6%; Score 856; DB 10; Length 1128;

Best Local Similarity 96.2%; Pred. No. 4.8e-217;		Matches 877; Conservative 0; Mismatches 35; Indels 0; Gaps 0;	
QY	4	AAACGTATTTATTTAAGAACTACCATCTGGTTGGATGGCCATTTGGGCAACC	63
Db	217	AAACGTATTTATTTAAGAACTACCATCTGGTTGGATGGCCATTTGGGCAACC	276
QY	64	TTTGACCTTACATCTGCCAAGCAATGTTCAATCCAAAGGTGCCATCTCACAAACAC	123
Db	277	TTTGACCTTACATCTGCCAAGCAATGTTCAATCCAAAGGTGCCATCTCACAAACAC	336
QY	124	CGCTCATCTGACAAATCCCATGGCTGCTGATTTACCATAGACATCAGCTGGAT	183
Db	337	CGCTCATCTGACAAATCCCATGGCTGCTGATTTACCATAGACATCAGCTGGAT	396
QY	184	CTGACTAACTTACCTCAGCAGGCGCCACCCCTTCAGAAATGGATTTGGATTA	243
Db	397	CTGACTAACTTACCTCAGCAGGCGCCACCCCTTCAGAAATGGATTTGGATTA	456
QY	244	GAGTCACCACTACACACCCCAAAAGAGTGGCATTTGACACTGTTCAACCTGACTTA	303
Db	457	GAGTCACCACTACACACCCCAAAAGAGTGGCATTTGACACTGTTCAACCTGACTTA	516
QY	304	ACTTATCGCGGTATTCAGATATCCAAAGTGGCTTATGCGCTTGGAGGACACAAAT	363
Db	517	ACTTATCGCGGTATTCAGATATCCAAAGTGGCTTATGCGCTTGGAGGACACAAAG	576
QY	364	CCCTTTGTGTTGAAGTGCACCAAGAGAGTGGTGTGCTGGTTGTGAGTAACCTGG	423
Db	577	CCCTTTGTGTTGAAGTGCACCAAGAGAGTGGTGTGCTGGTTGTGAGTAACCTGG	636
QY	424	AACCTTGAGCATCCAGGCTCAAGATATTACAGAGCTCAGCAAGATATTGAATCCAC	483
Db	637	AATCCGAGCATCCAGGCTCAAGATATTACAGAGCTCAGCAAGATATTGAATCCAC	696
QY	484	ACCTATGGCCAAACATTCGGAGAAATAGTGAAGATTTAAATCTGATCCCATCTCT	543
Db	697	ACCTATGGCCAAACATTCGGAGAAATAGTGAAGATTTAAATCTGATCCCATCTCT	756
QY	544	ACTTGTAAATTTATCTTTCATTTGAAAACATCAATTCACAAAGATTTACATCAGAGAAAG	603
Db	757	ACTTGTAAATTTATCTTTCATTTGAAAACATCAATTCACAAAGATTTACATCAGAGAAAG	816
QY	604	CTCTACAAATGATTTTGGCTGCTCAGTACTGTCTGCTGGGCTCCATCTAGGAAAC	663
Db	817	CTCTACAAATGATTTTGGCTGCTCAGTACTGTCTGCTGGGCTCCATCTAGGAAAC	876
QY	664	TATGAGAAATATTTATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC	723
Db	877	TATGAGAAATATTTATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC	936
QY	724	GAGTGAAGAAATATTCATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC	783
Db	937	GAGTGAAGAAATATTCATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC	996
QY	784	AACGTGAGAAAGATTTTACTGTAACCTACACAGGTTTGGGAATTCATTCATTCATTCATTC	843
Db	997	AACGTGAGAAAGATTTTACTGTAACCTACACAGGTTTGGGAATTCATTCATTCATTCATTC	1056
QY	844	GCATGGAGCAATGTAAGAGGATCAAGAAATATTAAGCTCTGTTGTAATTTAGAAATGG	903
Db	1057	GCATGGAGCAATGTAAGAGGATCAAGAAATATTAAGCTCTGTTGTAATTTAGAAATGG	1116
QY	904	TTTTGGAATTTAA 915	
Db	1117	TTTTGGAATTTAA 1128	
RESULT 3			
AB049819 2156 bp mRNA linear ROD 18-JAN-2002			
LOCUS Rattus norvegicus FUT9 mRNA for alpha1,3-fucosyltransferase IX,			
DEFINITION complete cds.			

ACCESSION	AB049819	GI:13591588
VERSION	AB049819.1	
KEYWORDS		
SOURCE	Rattus norvegicus	cDNA to mRNA.
ORGANISM	Rattus norvegicus	
REFERENCE	1 (sites)	
AUTHORS	Shimoda,Y., Tajima,Y., Oseanal,T., Katsune,A., Kohara,M., Kudo,T., Marimatsu,H., Takashima,N., Ishii,Y., Nakamura,S., Osumi,N. and Sanai,Y.	
TITLE	Pax6 Controls the Expression of Lewis x Epitope in the Embryonic Forebrain by Regulating alpha 1,3-Fucosyltransferase IX Expression	
JOURNAL	J. Biol. Chem. 277 (3), 2033-2039 (2002)	
PUBMED	11675393	
REFERENCE	2 (bases 1 to 2156)	
AUTHORS	Sanai,Y.	
TITLE	Direct Submision	
JOURNAL	Submitted (11-OCT-2000) Yutaka Sanai, Tokyo Metropolitan Institute of Medical Science, Department of Biochemical Cell Research, Honkomagome 3-18-22, Bunkyo-Ku, Tokyo 113-8613, Japan (E-mail:sanai@insboken.or.jp, Tel:81-3-3823-2101(ex.5233), Fax:81-3-3828-6663)	
FEATURES	Location/Qualifiers	
source	1..2156	
gene	/organism="Rattus norvegicus" /db_xref="taxon:10116" 346..1425 /gene="FUT9" 346..1425 /gene="FUT9" /codon_start=1 /product="alpha1,3-fucosyltransferase IX" /protein_id="BAB40953.1" /db_xref="GI:13591589" /translation="MTSTKSGILRPELIVCIILGCFVACLIIYIKPTNSVSPMESA SSVLKMKNFESRKTDFNETTILVWVPGSGTDFDLSCAMNIOGCHLTPTRSLYRK SHAVLIHNRDIEMDLNLPOLAPRPOKIMNMLESPHTPOKSGLEHLENTLYRK DSIDIOVPEGLTSTSPVEVPSKELICVGMNPNHARKYVNELEKSTIEHTY GOAFGEVYNDKMLIPRTISCKPLSPENSIRHODYTERIKYINAFIAGSVYVLPRESN YENYTRADSFTHVEDNPSSELAETLEVDKNNKLYLTFNMKDFVNLPRWESHA CLACDVKRHOEKRSVNLKMFWN"	
CDS		
BASE COUNT	605 a 470 c 437 g 644 t	
ORIGIN		
Query Match	93.2%; Score 852.8; DB 10; Length 2156;	
Best Local Similarity	95.9%; Pred. No. 3.4e-216;	
Matches 875; Conservative	0; Mismatches 37; Indels 0; Gaps 0;	
QY	4	AAACGTATTTATTTAAGAACTACCATCTGGTTGGATGGCCATTTGGGCAAGAC 63
Db	514	AAACGTATTTATTTAAGAACTACCATCTGGTTGGATGGCCATTTGGGCAAGAC 573
QY	64	TTTGACCTTACATCTGCCAAGCAATGTTCAATTCAAAGGTGCCATCTCACAAACAC 123
Db	574	TTTGACCTTACATCTGCCAAGCAATGTTCAATTCAAAGGTGCCATCTCACAAACAC 633
QY	124	CGCTCATCTGACAAATCCCATGGCTGCTGATTTACCATAGACATCAGCTGGAT 183
Db	634	CGCTCATCTGACAAATCCCATGGCTGCTGATTTACCATAGACATCAGCTGGAT 693
QY	184	CTGACTAACTTACCTCAGCAGGCGCCACCCCTTCAGAAATGGATTTGGATTA 243
Db	694	CTGACTAACTTACCTCAGCAGGCGCCACCCCTTCAGAAATGGATTTGGATTA 753
QY	244	GAGTCACCACTACACACCCCAAAAGAGTGGCATTTGACACTGTTCAACCTGACTTA 303
Db	754	GAGTCACCACTACACACCCCAAAAGAGTGGCATTTGACACTGTTCAACCTGACTTA 813
QY	304	ACTTATCGCGGTATTCAGATATCCAAAGTGGCTTATGCGCTTGGAGGACACAAAT 363
Db	814	ACTTATCGCGGTATTCAGATATCCAAAGTGGCTTATGCGCTTGGAGGACACAAAG 873

OY 364 CCCTTGTCTGTTGAAGTCCAGACGAGAGAGTGTGTGCTGGGTTGTGATTAAGTAC 423
|||||
Db 874 CCTTTGTGTGTTGAAGTCCAGACGAGAGAGTGTGTGCTGGGTTGTGATTAAGTAC 933
|||||
OY 424 AACCTGACATGACGAGGTCAGATTAACAGAGCTCAGCAAGATATTGAATCCAC 483
|||||
Db 934 AATCCGACATGACGAGGTCAGATTAACAGAGCTCAGCAAGATATTGAATCCAC 993
|||||
OY 484 ACCTATGACGAGATTCGAGATTAACAGATTAACAGATTAACAGATTAACAGATTA 543
|||||
Db 994 ACCTATGACGAGATTCGAGATTAACAGATTAACAGATTAACAGATTAACAGATTA 1053
|||||
OY 544 ACTTGTAAATTTATCTTCAATTTGAAACATTCACAAAGATTAACATTCACAAAG 603
|||||
Db 1054 ACTTGTAAATTTATCTTCAATTTGAAACATTCACAAAGATTAACATTCACAAAG 1113
|||||
OY 604 CTCTAATGACATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
|||||
Db 1114 CTCTAATGACATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1173
|||||
OY 664 TATGAGATTAATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 723
|||||
Db 1174 TATGAGATTAATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1233
|||||
OY 724 GACTTGAACAAATATCTGAGAGATTAACAAACATTAAGTTTACCTTGTACTT 783
|||||
Db 1234 GACTTGAACAAATATCTGAGAGATTAACAAACATTAAGTTTACCTTGTACTT 1293
|||||
OY 784 AACTGAGAAAGATTTTACCTGATTAACCTGATTAACCTGATTAACCTGATTAACCTG 843
|||||
Db 1294 AACTGAGAAAGATTTTACCTGATTAACCTGATTAACCTGATTAACCTGATTAACCTG 1353
|||||
OY 844 GCATGCGACCATGTAAGAAAGCATCAAGATTAATTAAGTCTGTTGTTGTTGTTGTTG 903
|||||
Db 1354 GCATGCGACCATGTAAGAAAGCATCAAGATTAATTAAGTCTGTTGTTGTTGTTGTTG 1413
|||||
OY 904 TTTTGGAAATTA 915
|||||
Db 1414 TTTTGGAAATTA 1425
|||||

RESULT 4
AF230460 1705 bp mRNA linear ROD 12-JUL-2000
LOCUS Cricetulus griseus alpha(1,3)fucosyltransferase type IX (Put9)
DEFINITION mRNA, complete cds.
ACCESSION AF230460
VERSION AF230460.1 GI:9049663
KEYWORDS
SOURCE Chinese hamster.
ORGANISM Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE 1 (bases 1 to 1705)
AUTHORS Patnaik,S.K., Zhang,A., Shi,S. and Stanley,P.
TITLE Alpha(1,3)fucosyltransferases expressed by the gain-of-function
Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30
JOURNAL Arch. Biochem. Biophys. 375 (2), 322-332 (2000)
MEDLINE 2016953
PubMed 10700388
REFERENCE 2 (bases 1 to 1705)
AUTHORS Patnaik,S.K., Shi,S. and Stanley,P.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of
Medicine, Chanln 516, 1300 Morris Park Avenue, New York, NY 10461,
USA

FEATURES
source location/Qualifiers
1..1705
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
1..1705
gene

CDS
/gene="Put9"
/note="CgPut9"
62..1141
/gene="Put9"
/note="Fuc-TIX"
/codon_start=1
/product="alpha(1,3)fucosyltransferase type IX"
/protein_id="AA82412.1"
/db_xref="GI:9049664"
/translation="MTSTSGILRPILVYVILACFMAFLIYIKPINSWFSPMESA
SSVLMKNEFSTKDYENFTTLLVWVGQGFDTSCQAFNIOGCHLTPRSLYRK
SHAYLHHRIISMDLNLPOQARPPQKIMWNLSPHPPKSGIEHLFNLITLYR
DSIDIOVFGILYTNDFEVPESKEKLYCWYKSWNDEARXYXNELEKSEIHTY
GOARVEYVNEKNIIPITISCKEYLSPEHSIKHDYITREKNAFLAGSVFVLAGPSEN
YENYIPADSPFIHEDYINSBELAKYILEYDKNNKILYISFNMKDFVNLPRWESHA
CLACDHRARHDEYKSGNLEKFWN"

BASE COUNT 515 a 336 c 320 g 534 t
ORIGIN

Query Match 90.2%; Score 825.4; DB 10; Length 1705;
Best Local Similarity 93.9%; Pred. No. 6,7e-209;
Matches 859; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

OY 1 ACAAACATGATTAATTTAAGAAATCAATCTGCTGGTGGTATGCGCATTTGGGAG 60
|||||
Db 227 ACAAACATGATTAATTTAAGAAATCAATCTGCTGGTGGTATGCGCATTTGGGAG 286
|||||
OY 61 ACCCTTGACCTTAACATCCGCGCAAGCATGTTCAATATCCAAAGGTGCCATCTCAACA 120
|||||
Db 287 ACCCTTGACCTTAACATCCGCGCAAGCATGTTCAATATCCAAAGGTGCCATCTCAACA 346
|||||
OY 121 GACCGCTCATTTGACAAATATCCATGCGGTCTGATTCACCATAGAGACATCAGCTGG 180
|||||
Db 347 GACCGCTCATTTGACAAATATCCATGCGGTCTGATTCACCATAGAGACATCAGCTGG 406
|||||
OY 181 GATCTGACTAATTTACTCTCAGACAGGAGCCAGCCCTTCACAAATGATTTGGATGAT 240
|||||
Db 407 GATCTGACTAATTTACTCTCAGACAGGAGCCAGCCCTTCACAAATGATTTGGATGAT 466
|||||
OY 241 TTAGAGTCAACCCAGCTACACCCGCCAAAGAGTGGCATGGAACATGTTGACCTGACT 300
|||||
Db 467 CTGAGTCTCCCATCATCATACCCGCCAAAGAGTGGCATGGAACATGTTGACCTGACT 526
|||||
OY 301 CTGAGTCTCCCATCATCATACCCGCCAAAGAGTGGCATGGAACATGTTGACCTGACT 360
|||||
Db 527 CTGAGTCTCCCATCATCATACCCGCCAAAGAGTGGCATGGAACATGTTGACCTGACT 586
|||||
OY 361 AATCCCTTGTGTTGAAGTCCAGACGAGAGAGTGTGCTGGGTTGTGAGTAAC 420
|||||
Db 587 AATCCCTTGTGTTGAAGTCCAGACGAGAGAGAGTGTGCTGGGTTGTGAGTAAC 646
|||||
OY 421 TGGAAACCTGAGCATGCGAGGTCAGATTAATTAACAGAGCTCAGCAAGATTAAGTAATC 480
|||||
Db 647 TGGAAACCTGAGCATGCGAGGTCAGATTAATTAACAGAGCTCAGCAAGATTAAGTAATC 706
|||||
OY 481 CACACCATGAGCGACATGCGAGGTCAGATTAATTAACAGAGCTCAGCAAGATTAAGTAATC 540
|||||
Db 707 CACACCATGAGCGACATGCGAGGTCAGATTAATTAACAGAGCTCAGCAAGATTAAGTAATC 766
|||||
OY 541 TCTACTGTTAATTTATCTTCAATTTGAAACATCAATTCACAAAGATTTACTACAGAA 600
|||||
Db 767 TCTACTGTTAATTTATCTTCAATTTGAAACATCAATTCACAAAGATTTACTACAGAA 826
|||||
OY 601 AAGCTTCAACATGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
|||||
Db 827 AAGCTTCAACATGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 886
|||||
OY 661 AACTATGAGATTAATATCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 720
|||||
Db 887 AACTATGAGATTAATATCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 946
|||||
OY 721 AAGTATGAGATTAATATCCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 780
|||||

Db 947 AGTACGTTGGCAAAATATCTTAAGAGTGTACAAAAACAATAAGTGTACTTACTTAT 1006
Qy 781 TTTAACTGGAGAAAGATTTTACTGTAAACCTACACAGGTTTGGAGATCATGATGC 840
Db 1007 TTTAACTGGAGAAAGATTTTACTGTAAACCTACACAGGTTTGGAGATCATGATGC 1066
Qy 841 CTGGCATGCCACATGTAAAGGATCAAGATATAGTCTGTGGTATTTAGAGAA 900
Db 1067 CTGGCTGTGTATCATCAATAAAGGACACAGAGATATAGTGTGGTATTTAGAGAA 1126
Qy 901 TGGTTTGAATTA 915
Db 1127 TGGTTTGAATTA 1141

RESULT 5
HSA238701 2501 bp mRNA linear PRI 21-OCT-2000
LOCUS HSA238701 Homo sapiens mRNA for alpha-3-fucosyltransferase.
DEFINITION AJ238701
ACCESSION AJ238701.1 GI:4741566
VERSION alpha-3-fucosyltransferase; FUT9 gene.
KEYWORDS SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2501)
Calleau-Thomas, A., Coullin, P., Candelier, J. J., Balazsino, L.,
Mennesson, R., Oriol, R. and Mollicone, R.
FUT9 and FUT9 genes are expressed early in human embryogenesis
Glycobiology 10 (8), 789-802 (2000)
2 (bases 1 to 2501)
Mollicone, R.
Direct Submission
Submitted (29-APR-1999) Mollicone R., Glycobiology, INSERM U504, 16
Av. Paul Vaillant-Couturier, Villejuif, 94807 Cedex, FRANCE
JOURNAL related sequence: AB015426.
COMMENT Location/Qualifiers
FEATURES
source 1..2501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/tissue_type="whole embryo"
/dev_stage="40 to 60 days"
105..2466
/gene="FUT9"
105..1184
/gene="FUT9"
/product="alpha-3-fucosyltransferase"
/codon_start=1
/protein_id="CAB41890.1"
/db_xref="GI:4741567"
/db_xref="SPTREMBL:Q9Y231"
/translation="MTSTKGLRPLVCIILGCFMACLIYIKPTNSWIFSPMESA
SSVLKMKNEFTKTDYFNETITLVWVWPGQTFDITSCQAMNIGCHLTUDRSLYNK
SHAVLIHHRDLSMDLNLPOQARPPFOKIMWNLESPTHTPOKSGIEHLENTLTYYR
DSDIQVPGFLSTNPFVEYSPKELICWVSNMNEPHARKYINELSKSEIETHY
GOAFGEYVNDKRLIPTISACKRYLSFENSIHDDYITELKYNAFLAGSPVYVPSREN
YENYIPADSEFIVEIDYNSPSELAKYLEVDKNKILYLSIFNKRKDFYVNLPRFWESHA
CLACDHVRHQDYKSVGNLEKWMN"
2192
/gene="FUT9"
/evidence="experimental"
2235
/gene="FUT9"
/evidence="experimental"
2310
/gene="FUT9"
/evidence="experimental"
2466
/gene="FUT9"
/evidence="experimental"

polyA_site
polyA_site
polyA_site
polyA_site

BASE COUNT 818 a 436 c 422 g 825 t
ORIGIN
Query Match 87.1%; Score 796.6; DB 9; Length 2501;
Best local similarity 91.9%; Pred. No. 3.1e-201;
Matches 841; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
Qy 1 ACAAAGTGTATTTTAAAGAACTACCATCTGTGGTGGTATGGCATTTGGGAG 60
Db 270 ACCAAAGTGTATTTTAAAGAACTACCATCTGTGGTGGTATGGGAG 329
Qy 61 ACCCTTGACCTTACATCCGCGCAAGCAATGTATATCCAGAGGTCATCTCACAA 120
Db 330 ACCCTTGACCTTACATCCGCGCAAGCAATGTATATCCAGAGGTCATCTCACAA 389
Qy 121 GACCGCTATGTACAAACAAATCCCATGGCGGCTCTGATTCACATGAGACATCAGCTG 180
Db 390 GACCGCTATGTACAAACAAATCCCATGGCGGCTCTGATTCACATGAGACATCAGCTG 449
Qy 181 GATCTGACTTAACCTTACCTCAGCAGGCGCAGCCACCTTCAGAAATGATTTGGATGA 240
Db 450 GATCTGACTTAACCTTACCTCAGCAGGCGCAGCCACCTTCAGAAATGATTTGGATGA 509
Qy 241 TTAGAGTACCCACTACACACCCCAAAAGAGTGCGCATTTGAACCTTGTCACTGACT 300
Db 510 TTGGAATCACCACTACACACTCCCAAAAGAGTGCGCATTTGAACCTGACT 569
Qy 301 CTAACTTACGCGCGGTATTCAGATTCAGAGTCCCTTATGGCTCTTGAACGGTGACAG 360
Db 570 CTAACTTACGCGCGGTATTCAGATTCAGAGTCCCTTATGGCTCTTGAACGGTGACAG 629
Qy 361 AATCCCTTGTGTGTTGAAGTGCCAAAGCAAGAGAGTGTGCTGCGTGTGAGTAAC 420
Db 630 AATCCCTTGTGTGTTGAAGTGCCAAAGCAAGAGAGTGTGCTGCGTGTGAGTAAC 689
Qy 421 TGGAACTCTGAGATGCGCAGGCTCAAGATTTACACAGAGCTCAGCAAGATTTGAATC 480
Db 690 TGGAACTCTGAGATGCGCAGGCTCAAGATTTACAAAGAGCTCAGCAAGATTTGAATC 749
Qy 481 CACACCTATGGCCAGCATTCGAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 540
Db 750 CACACCTATGGCCAGCATTCGAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 809
Qy 541 TCTACTTGTAAATTTATCTTTCATTTGAAACCTCAATTCACAAAGATTTACACAGAA 600
Db 810 TCTACTTGTAAATTTATCTTTCATTTGAAATTCATTCACAAAGATTTACACAGAA 869
Qy 601 AAGCTTACAAATGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 870 AAGCTTACAAATGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
Qy 661 AACTATGAGAAATTAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 720
Db 930 AACTATGAGAAATTAATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 989
Qy 721 AGTGAATGCAAAATATCTGAAAGAGTGAACAAACAAATTAAGTTGACCTTACTTAC 780
Db 990 AGTGAATGCAAAATATCTGAAAGAGTGAACAAACAAATTAAGTTAAGTTACTTACT 1049
Qy 781 TTTAACTGGAGAAAGATTTTACTGTAAACCTACACAGGTTTGGAGATTCATGATGC 840
Db 1050 TTTAACTGGAGAAAGATTTTACTGTAAACCTACACAGGTTTGGAGATTCATGATGC 1109
Qy 841 CTGGCATGCCACATGTAAAGGATCAAGAAATTAAGTCTGTGGTATTTAGAGAA 900
Db 1110 CTGGCATGCCACATGTAAAGGATCAAGAAATTAAGTCTGTGGTATTTAGAGAA 1169
Qy 901 TGGTTTGAATTA 915
Db 1170 TGGTTTGAATTA 1184

RESULT 6

AL512406 AL512406 80247 bp DNA linear PRI 27-APR-2001
LOCUS Human DNA sequence from clone RP11-504J9 on chromosome 6, complete
DEFINITION sequence.
ACCESSION AL512406
VERSION AL512406
KEYWORDS HTG.
SOURCE HTG.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 80247)
Bates,K.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Apr 30, 2001 this sequence version replaced g1:13396709.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL, SW:
SWISSPROT, Tr: TrEMBL, Wp: WormPep; Information on the WormPep
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormPep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-504J9 is from the library RP11-77B15 is at 100 in this sequence.
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBac3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-504J9. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP1-381A2 is at 80148 in this sequence.
The true right end of clone RP11-77B15 is at 100 in this sequence.
location/Qualifiers
source 1. 80247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-504J9"
/clone_1kb="RP11-11.2"
184. 1422
repeat_region /note="12 repeat: matches 875. .2419 of consensus"
1744. 1834
repeat_region /note="MIR repeat: matches 166. .258 of consensus"
2950. 2977
repeat_region /note="14 copies 2 mer ca 100% conserved"
3411. 3669
repeat_region /note="LTR16C repeat: matches 93. .364 of consensus"
4372. 4704
repeat_region /note="LTR1A1 repeat: matches 5. .365 of consensus"
complement(7821..8406)
misc_feature /note="match: GSS: Em:A0550748"
8126. 8184
repeat_region /note="MER58B repeat: matches 272. .332 of consensus"
8185. 8696
repeat_region /note="MUT2D repeat: matches 1. .513 of consensus"
8779. 8817
/note="MUT2D repeat: matches 504. .542 of consensus"
8819. .8974
repeat_region /note="MER58A repeat: matches 6. .162 of consensus"
complement(9443..9891)
misc_feature /note="match: GSS: Em:B90633"
complement(9452..9883)
misc_feature /note="match: GSS: Em:AQ119695"
9782. 9894
repeat_region /note="MIR repeat: matches 92. .205 of consensus"
10089. 10212
repeat_region /note="MIR repeat: matches 135. .262 of consensus"
complement(10093..10671)
misc_feature /note="match: GSS: Em:AQ237061"
10223. 10282
repeat_region /note="MIR repeat: matches 203. .262 of consensus"
complement(10315..10631)
misc_feature /note="match: GSS: Em:AQ200259"
10805. 11186
repeat_region /note="match: GSS: Em:AQ072353"
11657. 11844
repeat_region /note="12 repeat: matches 2546. .2750 of consensus"
complement(11659..12084)
misc_feature /note="match: GSS: Em:A0518361"
complement(11674..11967)
misc_feature /note="match: GSS: Em:AQ034678"
12083. 12588
repeat_region /note="match: GSS: Em:AQ691235"
12596. 12810
repeat_region /note="MIR repeat: matches 23. .248 of consensus"
12920. 12986
repeat_region /note="L1MA5 repeat: matches 6228. .6295 of consensus"
12987. 13169
repeat_region /note="L1MA5 repeat: matches 5852. .6043 of consensus"
13168. 13280
repeat_region /note="L1MA8 repeat: matches 5710. .5832 of consensus"
13851. 17218
repeat_region /note="L1P1 repeat: matches 2787. .6155 of consensus"
18903. 18953
repeat_region /note="MUT2CB repeat: matches 439. .501 of consensus"
18988. 19433
repeat_region /note="MUT2CB repeat: matches 1. .460 of consensus"
complement(19108..19690)
misc_feature /note="match: GSS: Em:A053935"
20683. 20801
repeat_region /note="AluYo/FRAM repeat: matches 169. .288 of consensus"
20802. 20862
repeat_region /note="AluYo/FRAM repeat: matches 3. .66 of consensus"
20912. 21033
repeat_region /note="MIR repeat: matches 114. .237 of consensus"
21047. 21128
repeat_region /note="41 copies 2 mer ta 63% conserved"
21676. 21859
repeat_region /note="12 repeat: matches 2130. .2308 of consensus"
21890. 23612
repeat_region /note="L1P repeat: matches 841. .2570 of consensus"
23601. 26609
repeat_region /note="L1PA5 repeat: matches 3107. .6145 of consensus"
26620. 26789
repeat_region /note="MER5A repeat: matches 1. .174 of consensus"
28521. 28627
repeat_region /note="MER39 repeat: matches 5. .111 of consensus"
28656. 28837
repeat_region /note="MER39B repeat: matches 54. .235 of consensus"
28838. 29142
repeat_region /note="AluSc repeat: matches 4. .308 of consensus"
29143. 29494
repeat_region /note="MER39B repeat: matches 235. .549 of consensus"
30529. 31344
repeat_region /note="L1PA3 repeat: matches 5331. .6146 of consensus"
32234. 32340
repeat_region /note="12 repeat: matches 2627. .2750 of consensus"
32399. 32578
repeat_region /note="MIR repeat: matches 12. .203 of consensus"

Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15932
Center clone name: 167_K_24

* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
789 888: contig of 788 bp in length
889 889: gap of 100 bp
1703 1802: contig of 814 bp in length
1803 2642: contig of 840 bp in length
2643 2742: gap of 100 bp
2743 3637: contig of 895 bp in length
3638 3737: gap of 100 bp
3738 4635: contig of 898 bp in length
4636 4735: gap of 100 bp
4736 5657: contig of 922 bp in length
5658 5757: gap of 100 bp
5758 6692: contig of 935 bp in length
6693 6792: gap of 100 bp
6793 7703: contig of 911 bp in length
7704 7803: gap of 100 bp
7804 8742: contig of 939 bp in length
8743 8842: gap of 100 bp
8843 9787: contig of 945 bp in length
9788 9887: gap of 100 bp
9888 10819: contig of 932 bp in length
10820 10919: gap of 100 bp
10920 11808: contig of 889 bp in length
11809 11908: gap of 100 bp
11909 12830: contig of 922 bp in length
12831 12930: gap of 100 bp
12931 13847: contig of 917 bp in length
13848 13947: gap of 100 bp
13948 14886: contig of 939 bp in length
14887 14986: gap of 100 bp
14987 15911: contig of 925 bp in length
15912 16011: gap of 100 bp
16012 16903: contig of 892 bp in length
16904 17003: gap of 100 bp
17004 17906: contig of 903 bp in length
17907 18006: gap of 100 bp
18007 18931: contig of 925 bp in length
18932 19031: gap of 100 bp
19032 19945: contig of 914 bp in length
19946 20045: gap of 100 bp
20046 20951: contig of 906 bp in length
20952 21051: gap of 100 bp
21052 21949: contig of 898 bp in length
21950 22049: gap of 100 bp
22050 22950: contig of 901 bp in length
22951 23050: gap of 100 bp
23051 23945: contig of 895 bp in length
23946 24045: gap of 100 bp
24046 24955: contig of 910 bp in length
24956 25055: gap of 100 bp
25056 25980: contig of 925 bp in length
25981 26080: gap of 100 bp
26081 27052: contig of 972 bp in length
27053 27152: gap of 100 bp
27153 28067: contig of 915 bp in length
28068 28167: gap of 100 bp
28168 29076: contig of 909 bp in length

29077 29176: gap of 100 bp
29177 30093: contig of 917 bp in length
30094 30193: gap of 100 bp
30194 31099: contig of 906 bp in length
31099 31199: gap of 100 bp
31200 32097: contig of 896 bp in length
32098 32197: gap of 100 bp
32198 33113: contig of 916 bp in length
33114 33213: gap of 100 bp
33214 34125: contig of 912 bp in length
34126 34225: gap of 100 bp
34226 35133: contig of 908 bp in length
35134 35233: gap of 100 bp
35234 36150: contig of 917 bp in length
36151 36250: gap of 100 bp
36251 37171: contig of 921 bp in length
37172 37271: gap of 100 bp
37272 38159: contig of 888 bp in length
38160 38259: gap of 100 bp
38260 39156: contig of 897 bp in length
39157 39256: gap of 100 bp
39257 40188: contig of 933 bp in length
40189 40288: gap of 100 bp
40289 41210: contig of 922 bp in length
41211 41310: gap of 100 bp
41311 42222: contig of 912 bp in length
42223 42322: gap of 100 bp
42323 43218: contig of 896 bp in length
43219 43318: gap of 100 bp
43319 44233: contig of 915 bp in length
44234 44333: gap of 100 bp
44344 45246: contig of 913 bp in length
45247 45346: gap of 100 bp
45347 46261: contig of 915 bp in length
46262 46361: gap of 100 bp
46362 47279: contig of 918 bp in length
47280 47379: gap of 100 bp
47380 48276: contig of 897 bp in length
48277 48376: gap of 100 bp
48377 49282: contig of 906 bp in length
49283 49382: gap of 100 bp
49383 50305: contig of 923 bp in length
50306 50405: gap of 100 bp
50406 51314: contig of 909 bp in length
51315 51414: gap of 100 bp
51415 52335: contig of 921 bp in length
52336 52435: gap of 100 bp
52436 53163: contig of 728 bp in length
53164 53263: gap of 100 bp
53264 54210: contig of 947 bp in length
54211 54310: gap of 100 bp
54311 55199: contig of 889 bp in length
55200 55299: gap of 100 bp
55300 56212: contig of 913 bp in length
56213 56312: gap of 100 bp
56313 57287: contig of 907 bp in length
57288 57387: gap of 100 bp
57388 58245: contig of 856 bp in length
58246 58345: gap of 100 bp
58346 59236: contig of 891 bp in length
59237 59336: gap of 100 bp
59337 60272: contig of 936 bp in length
60273 60372: gap of 100 bp
60373 61289: contig of 917 bp in length
61290 61389: gap of 100 bp
61390 62299: contig of 910 bp in length
62400 62399: gap of 100 bp
62400 63322: contig of 923 bp in length
63323 63422: gap of 100 bp
63423 64336: contig of 914 bp in length
64337 64436: gap of 100 bp
64437 65354: contig of 918 bp in length
65355 65454: gap of 100 bp

* 65455 66352: contig of 898 bp in length
 * 66353 66452: gap of 100 bp
 * 66453 67362: contig of 910 bp in length
 * 67363 67462: gap of 100 bp
 * 67463 68370: contig of 908 bp in length
 * 68371 68470: gap of 100 bp
 * 68471 68479: contig of 909 bp in length
 * 68479 69479: gap of 100 bp

Query Match 73.0%; Score 668.4; DB 2; Length 71396;
 Best Local Similarity 83.2%; Pred. No. 4.2e-167;
 Matches 764; Conservative 0; Mismatches 140; Indels 14; Gaps 6;

1 ACMAAACTGATTATTTTACGAAACTACCATTCCTGGTTGGGTATGGCCATTGGGCGAG 60
 Db 65550 ACMAAACTGATTATTTTACGAAACTACCATTCCTGGTTGGGTATGGCCATTGGGCGAG 65609
 Oy 61 ACCTTGACCTTACATCCTGCCAAGCAATGTTCAATATCCAGGGTGGCATCTCACACA 120
 Db 65610 ACCTTGACCTTACATCCTGCCAAGCAATGTTCAATATCCAGGGTGGCATCTCACACA 65669
 Oy 121 GACCGCTCATTTGTACAAACAAATCCAGCGCTCCATATCCAGATAGACATAGAGCTGG 180
 Db 65670 GACCGCTCATTTGTACAAACAAATCCAGCGCTCCATATCCAGATAGACATAGAGCTGG 65729
 Oy 181 GATCTGACTAATCTTACCTCAGACAGGCGCACCCCTTCAGAAATGATTTGGATGAAT 240
 Db 65730 GATCTGACTAATCTTACCTCAGACAGGCGCACCCCTTCAGAAATGATTTGGATGAAT 65789
 Oy 241 TTAGAGTACCCACTCACACCCCAAAAGATGGCATTTGAACACTTTTCACCTGACT 300
 Db 65790 TTAGAGTACCCACTCACACCCCAAAAGATGGCATTTGAACACTTTTCACCTGACT 65849
 Oy 301 CTAACTTATGCGCGTATTCAGATATCCAGATGCTTATGCGTTTTCAGCGTGAGCACA 360
 Db 65850 CTAACTTATGCGCGTATTCAGATATCCAGATGCTTATGCGTTTTCAGCGTGAGCACA 65909
 Oy 361 AATCCCTTTGTTTGAAGTGCAGAAAGAGAAATGTTGGTGGGTTTGAGATAC 420
 Db 65910 AATCCCTTTGTTTGAAGTGCAGAAAGAGAAATGTTGGTGGGTTTGAGATAC 65969
 Oy 421 TGGAACTTGAGCATGCGAAGGTGCAATATTAACAGCACTGACAGAGATTTGAAATC 480
 Db 65970 TGGAACTTGAGCATGCGAAGGTGCAATATTAACAGCACTGACAGAGATTTGAAATC 66029
 Oy 481 CACACCTATGCCAAGCAATTCGGAATACGTAAGCAATTAATCTGATTCGACACATA 540
 Db 66030 CACACCTATGCCAAGCAATTCGGAATACGTAAGCAATTAATCTGATTCGACACATA 66089
 Oy 541 TCTACTTGTAAATTTTATCTTCAATTTGAAACTCAATTCACAAAGATTATACACAGAA 600
 Db 66090 TCTACTTGTAAATTTTATCTTCAATTTGAAACTCAATTCACAAAGATTATACACAGAA 66146
 Oy 601 AAGCTCTACAAATGATTTTGGCTGCTGAGTACCTGTTGCTGCTGGGTCATATAGGAA 660
 Db 66147 AAGCTCTACAAATGATTTTGGCTGCTGAGTACCTGTTGCTGCTGGGTCATATAGGAA 66200
 Oy 661 AACTATGGAATTAATATATCCAGCTGATTCATTCATTCGATGATTCATTCATTCGCTC 717
 Db 66201 AACTATGGAATTAATATATCCAGCTGATTCATTCATTCGATGATTCATTCATTCGCTC 66260
 Oy 718 CCCAGTGAATTAATATATTCGAGAAATGATTCAGAAACAAATTAATGTT- G1ACTTAG 776
 Db 66321 CCCAGTGAATTAATATATTCGAGAAATGATTCAGAAACAAATTAATGTTG1ACTTAG 66320
 Oy 777 TTACTTTAACT- GAGAAAGGATTTTACTGTAACCTACACAGGTTTGGGAATCACATG 835
 Db 66321 TTACTTTAACTGAGAAAGGATTTTACTGTAACCTACACAGGTTTGGGAATCACATG 66380
 Oy 836 CATGCGTGGCATGGACCATGTAAAGGACATCAAGATATTAAGCTGTGTGTAATTAG 895
 Db 66381 NNN 66440

Oy 896 AGAAATGCTTTGGAAT 913
 Db 66441 NNN 66458

RESULT 9
 AB035906 1707 bp DNA linear VRT 05-JAN-2002
 LOCUS
 DEFINITION Gallus gallus gene for cfu9, complete cds.
 ACCESSION AB035906
 VERSION AB035906.1 GI:18146865
 KEYWORDS cfu9.
 SOURCE Gallus gallus
 ORGANISM Gallus gallus brain DNA.

REFERENCE
 AUTHORS Kaneko, M., Nishihara, S., Kitano, T., Narimatsu, H. and Saitou, N.
 TITLE The evolutionary history of glycosyltransferase genes
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 1707)
 AUTHORS Kaneko, M., Saitou, N. and Kitano, T.
 TITLE Direct Submission
 JOURNAL Submitted (17-DEC-1999) Muka Kaneko, National Institute of Genetics, Laboratory of Evolutionary Genetics, Yata 111, Mishima, Shizuoka 411-8540, Japan (E-mail: mkaneko@med.ri.yamagata-u.ac.jp, Tel: 81-559-81-6790, Fax: 81-559-81-6789)
 COMMENT Sequence updated (29-Feb-2000).
 FEATURES
 location/Qualifiers
 source 1..1707
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /feature_type="brain"
 gene 268..1347
 /gene="cfu9-11x"
 CDS 268..1347
 /gene="cfu9-11x"
 /codon_start=1
 /product="cfu9"
 /protein_id="BA82489.1"
 /db_xref="GI:18146865"
 /translation="MTSTKGIIFRPFIVFIVLCFMAFLIYIKPSTSMISPVESA
 SSYKMSSEFSSKTDNFNETILLVLMFPGQFELTSCQFMFIHCHLFTIDRLYRK
 SHAVLIHHRDINMDLNLPEOAPRPPOKIMWNLDESPHTPOKSGIENHLEFNTLYRR
 DSDIOPVGSIMIVTGAFYEVPSKENVKVMVSNMDEHARVYENELSKYIEITHY
 GOAGDYVNDKRLIPYISSCKEYLFSESIHKDYITEKLYNALAGSVYVLPSPREN
 YENIIPADSEIHVEDFLSPRELAELVLMIDKNNMYLSTYNNMRKDFSVILPFWESHA
 CIACDHYKRHOEYKSGINLEKFFVN"

BASE COUNT 517 a 329 g 521 t 2 others
 ORIGIN

Query Match 68.8%; Score 629.2; DB 5; Length 1707;
 Best Local Similarity 80.5%; Pred. No. 1.1e-156;
 Matches 736; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Oy 2 CAAAACGATATTTTACGAAACTACCATTCCTGGTTGGGTATGGCCATTGGGCGAGA 61
 Db 434 CAAAACGATATTTTATATATATGAAATGATTTGTTGGCTTGGCCATTGGGCGAGA 493
 Oy 62 CTTTGACCTTACATCTGCCAAGCAATGTTCAATATCCAAAGGTGACATCTCACACAG 121
 Db 494 CATTGATCTTAACATCTGCCAAGCAATGTTCAATATCCAAAGGTGACATCTGACTATG 553
 Oy 122 ACCGCTCATTTGACAAACAAATCCATGCGTCCGATTCACCATGAGACATGAGCTGGG 181
 Db 554 ACCGCTCATTTATATACAAAGTCCATGCTCATATCATCATGAGGACATTAATGCGG 613
 Oy 182 ATCTGACTAATCTTACATGAGGCGAGGCGACCCCTTCAGAAATGATTTGGATGAAT 241
 Db 614 ATCTGACTAATTTACTTACGACCAAGCGACACATTTCCAGAAATGATTTGGATGAAT 673
 Oy 242 TAGAGTACCCACTCACACCCCAAAAGATGGCATTTGAACACTTTGTTCAACTGACTC 301


```

Db 674 TGGATATCCCAATCATCTCCACAAAAGATGGCATTTGAACACCTCTTTAACTAACCC 733
OY 302 TAACTATCCGCGGATTCAGATATCCAGTGCCTTATGCGTTTGTACGGTGGACACA 361
Db 734 TGACTTACCGCGGATTCAGATATTCAGTGCCTTATGCGTTTGTACGGTGGACAG 793
OY 362 ATCCCTTGTGTTGAAGTCCACAAAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGT 421
Db 794 GCGCCTTATCATATGAGTCCAAAGTAAAGAAATTTGGTTGTGTGTGTGTGTGTGTGT 853
OY 422 GGAACCTTGAGCATTCGAGGTCAGATATTCACAGCTCAGCAGATATTTGAATTC 481
Db 854 GGAACCTTGAGCATTCGAGGTCAGATATTCACAGCTCAGCAGATATTTGAATTC 913
OY 482 ACACCTATGSCCAAGCATTCGAGGATACGTGAACGATTAATAATTCGATCCACCAT 541
Db 914 ATACCTATGSCCAAGCATTCGAGGATACGTGAACGATTAATAATTCGATCCACCAT 973
OY 542 CTACTGTATATTTTATCTTTCATTTGAAAACTCAATTCACAAAGATTTACATCAGAAA 601
Db 974 CCACTTGCACAAATTCATACCTTTTCAATTTGAAAAATTCACAAAGATTTACATCAGAAA 1033
OY 602 AGCTCTACATGCAATTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
Db 1034 AACCTTACAAATGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1093
OY 662 ACTATGGAATTTATATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 721
Db 1094 ACTATGGAATTTATATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1153
OY 722 GTGAGTTAGCAAAATATCTGAGAGAGTTCACAAAACAAATAGTTTACCTTAGTTACT 781
Db 1154 GAGAGCTTGGAGAAATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1213
OY 782 TTAACTTGGAAGAGATTTTACCTGTAACTTACAGGTTTGGGAATTCATGATGATCC 841
Db 1214 TCACCTGAGAGAGATTTTACCTGTAACTTACAGGTTTGGGAATTCATGATGATCC 1273
OY 842 TGGCATTGCGACATGTTTAAAGAGCATCAAGATATTAAGTGTGTGTGTGTGTGTGTGT 901
Db 1274 TTGCTTGTGATCATGTGAAAAAGACACAGAAATTAAGTGTGTGTGTGTGTGTGTGT 1333
OY 902 GGTTTGGGAATTA 915
Db 1334 GGTTTGGGAATTA 1347

RESULT 10
AB035905 1080 bp DNA linear VRT 05-JAN-2002
LOCUS xenopus laevis gene for xFUT9, complete cds.
DEFINITION AB035905
ACCESSION AB035905
VERSION AB035905.1 GI:18146863
KEYWORDS xFUT9.
SOURCE xenopus laevis DNA.
ORGANISM xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
1 (sites)
REFERENCE
AUTHORS Kaneko,M., Nishihara,S., Kitano,T., Nishimatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL Unpublished
AUTHORS Kaneko,M., Saitou,N. and Kitano,T.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1999) Miki Kaneko, National Institute of
Genetics, Laboratory of Evolutionary Genetics, Yata 111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6789)
COMMENT
FEATURES
Location/Qualifiers

```

```

source 1. 1080
/organism="Xenopus laevis"
/gene
CDS 1. 1080
/gene="xFuc-TIX"
1. 1080
/gene="xFuc-TIX"
/product="xFUT9"
/codon_start=1
/translation="MTSASKGLIRAFVAVSVLICFMICLLIVKPNMNISSPIESA
HSLKMKNEFSSKONYNETVLIWVPEFGFEKSCOTLENIYCHITTRALYNK
SHAVLHHHDIEMDLINLPOLRPPPOKIMNLESPTPTPOKSGIEHLENLTLYR
DSDIQVYGFENVSTKPPDFEVPKSKFVCVMVSNMNPBARVKNYELIKYIETTY
GGAPEVLSDKSLPLTSLASCKFYLSPENIHNDYIEXKYNALLAGSVIYIGPREN
YENYIPADSFIEHVEDPLSLRELDHLLINKDTBOYLTSTNNRKHFTVNMRFWESHA
CLACDHVRKHQEKSVGNLEKMWNVN"
BASE COUNT 332 a 231 c 202 g 315 t
ORIGIN
Query Match 64.4%; Score 589.2; DB 5; Length 1080;
Best Local Similarity 77.8%; Pred. No. 4,8e-146;
Matches 711; Conservative 0; Mismatches 203; Indels 0; Gaps 0;
OY 2 CAAAACGTATTTATTTTAAAGAACTACATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 61
Db 167 CTAAACAGATTTATTCATTAAGAAACCTGCTTCAATTTGGGTGTGTGTGTGTGTGTGT 226
OY 62 CTTTGCATTTACATTCCTGCAAGCAATGTTCATATTCACAAAGGTGTCATTCACACAG 121
Db 227 CTTTGCATTTACATTCCTGCAAGCAATGTTCATATTCACAAAGGTGTCATTCACACAG 286
OY 122 ACCGTCATGTTCACAAATTCACAAATTCACAAATTCACAAATTCACAAATTCACAAAT 181
Db 287 ACCGTCATGTTCACAAATTCACAAATTCACAAATTCACAAATTCACAAATTCACAAAT 346
OY 182 ATCTGACTTATTCCTGACAGGCGACACCTTTCAGAAATGATTTGGATTAAT 241
Db 347 ATCTGACTTATTCCTGACAGGCGACACCTTTCAGAAATGATTTGGATTAAT 406
OY 242 TAGAGTACCCACTCACACCCCCCAAAAAGATGAGCATTTGAACCTGTTCAACCTGATC 301
Db 407 TGGAAATCACCAACACATACGCCCAAAAAGATGAGCATTTGAACCTGTTCAACCTGATC 466
OY 302 TAACCTATCCCGCTGATTCAGATATTCAGATATTCAGATATTCAGATATTCAGATATTC 361
Db 467 TGACTTACAGCGGTGACTCAGATATTCAGATATTCAGATATTCAGATATTCAGATATTC 526
OY 362 ATCCCTTGTGTTGAAGTCCCAAGCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421
Db 527 AACCATTTGACTTTGAGGTGCCAAGCAAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 586
OY 422 GGAACCTTGAGCATTCGAGGTCAGATATTCACAGCTCAGCAGATATTTGAATTC 481
Db 587 GGAACCTTGAGCATTCGAGGTCAGATATTCACAGCTCAGCAGATATTTGAATTC 646
OY 482 ACACCTATGSCCAAGCATTCGAGGATACGTGAACGATTAATAATTCGATCCACCAT 541
Db 647 TAACCTATGSCCAAGCATTCGAGGATACGTGAACGATTAATAATTCGATCCACCAT 706
OY 542 CTACTGTATATTTTATCTTTCATTTGAAAACTCAATTCACAAAGATTTACATCAGAAA 601
Db 707 CACTTGTCAAAATTTTATCTTTCATTTGAAAACTCAATTCACAAAGATTTACATCAGAAA 766
OY 602 AGCTCTACATGCAATTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
Db 767 AACCTTATATGATATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 826
OY 662 ACTATGGAATTTATATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 721
Db 827 ACTATGGAATTTATATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTC 886

```

QY	722	GTGAGTGTACCAAAATATCTGAAAGAGTGTGACAAAAGAAACATTAAGTGTACCTGAGTACT	781
Db	887	GAGAATTACCTGACCATCTTTTATCTGACAAAAGATACGAGACGTCACCTGAGCTACT	946
QY	782	TTTATCTGAGAAAGATTTTTTACTGTAAACCTACACGAGTTTGGGAATCAACATGATGCC	841
Db	947	ACAACTGGAGGAACATTTTACAGTCGATATATGCTCCACTCTCTGGGAATCCCATGCTACT	1006
QY	842	TGGCAGCGACCACTGTAAAGGCACTCAAGAAATATAGTCTGTGGTAAATTTAGAGAAAT	901
Db	1007	TAGCGTGTACCACTGTAAAGGCACTCAAGAAATACAAATCAATGCTGGCAATCTAGAGAGT	1066
QY	902	GGTTTGTGGAAATTA	915
Db	1067	GGTCTTGCACTTA	1080

RESULT 11	AB023627	1567 bp	DNA	linear	VRT 02-APR-1999
LOCUS	AB023627				
DEFINITION	AB023627	Danio rerio zfm1 gene for alpha(1,3)fucosyltransferase, complete cds.			
ACCESSION	AB023627				
VERSION	AB023627.1	GI:4587295			
KEYWORDS	alpha(1,3)fucosyltransferase.				
SOURCE	Danio rerio DNA.				
ORGANISM	Danio rerio				

REFERENCE
1 (sites)

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
Molecular cloning and characterization of two zebrafish alpha(1,3)fucosyltransferase genes developmentally regulated in embryogenesis	J. Biochem. (1999)	In press	2 (bases 1 to 1567)		
			Kageyama, N., Natsuka, S. and Hase, S.		
Submitted (12-FEB-1999)			Shunji Natsuka, Osaka University, Graduate		

School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034, Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, Tel: 81-6-6850-5381, Fax: 81-6-6850-5383)

FEATURES	Location/Qualifiers
source	1. .1567

```

/organism="Danio rerio"
/db_xref="taxon:7955"
177. .1352
/gene="zfrr1"
177. .1352
CDS
177. .1352

```

```

//gene: zfit1
//codon_start:-1
//product:"alpha(1,3)fucosyltransferase"
//protein_id:"BA66706.1"
//db_xref:"GI:4587296"
//translation:"MDKTTLPSPSKAAQVYIATFMILSPVCIEFYVYVNPNTFEKFE
AVDHVNHSCSTFCGLDVILAKONKCTITKASNPSTPAPAPAPAAKADDDETILL
MYTFEAFRDLVCGSSQFNLIHGCHLIDDRSMQKAGVGFHHRDLASDI,PPPPRAFE
KILMMNNPSSNSHPNSLIYNGFNLTSSRKSDI,IPVYGGIL,IDADEQKNTTIPKIKK
KLVCMVSNFQEHKRSQVYVYVYELKHIKVEAAGVYVYVYVYVYVYVYVYVYVYVYVY
KSSNIHDEQYTFEFLPMLGIVTVPVGLSPGRDYVEEFLPRAFLVHDFPTPEMLADHII
KSDIHDEQYTFEFLPMLGIVTVPVGLSPGRDYVEEFLPRAFLVHDFPTPEMLADHII
KSSIHDEQYTFEFLPMLGIVTVPVGLSPGRDYVEEFLPRAFLVHDFPTPEMLADHII
334 c 293 g 441 t

```

Query Match	29.3%	Score 268.2;	DB 5;	Length 1567;
Best local Similarity	57.8%;	Pred. No. 1.3e-60;		
Matches 521; Conservative	0;	Mismatches 368;	Indels 12;	Gaps 2;

OY 18 TAACGAACCTACCATTCTGGTTGGGTATGCGCATTTGGCAGACCTTGCACCTTACATC 77
- - - - -
Db 461 TCAGAACAACATCTCTGTGATCTGGGTATCTCTTCGAGACGATTTTGACTTTGGGAGT 520

OY	78	CTGGACACATGTTCAATATCCAGGGGTGCATCTCACACAGACGGCTCATTTGACAA	137
Db	521	GTGTGGTTTGCAGTTCCACATCTCCAGCGCTGTCAATTTAACGACGACAGAAATATGATCCA	580
OY	138	CAATTCCTCATCGGTCCGTGATTCCACATPAGACATACGCTGGATCTGACTAATTACC	197
Db	581	GAACACGATGGAATTATGTTATGTTACCAAGAGATCTGAACGGGGAT-----TTGCC	631
OY	198	TCAGAGGCGCAGGCCACCCCTTTCAGAAATGGAATTTGATGAAATTTAGAGTCACCCACTCA	257
Db	632	GCACACCCCTTGACACAGCTTTCCAAAATGGATATGGTGAATATGAGTCTCCGAGTAA	691
OY	258	CACCCCCCAAAAGAGTGGCTTGAACACTTTGTTAAACCTGACCTTACTATATGCGCGTGA	317
Db	692	TTCCCATCCAAACAGCTGTGATTAATGATGSGTTTAATCTGACCTGCAATTTCCGCAAGGA	751
OY	318	TTTCAATATTCAGATGGCTTA---TGGCTTTTCACGGTGACACACAATCCCTTTGTGTT	374
Db	752	TTTCAATATTCAGATGGCTTACGGTGTGACTATGATGACACCGAAGCAAGAAAATT	811
OY	375	TCAGATGCCAAGCAGAGAGAGTTGGTGGCTGGTGTGAGTAACTGGAACCCCTGACGA	434
Db	812	CACACTCCCAAGAAAGGATTAACCTAGTTTCCGTGATAGTAACTTTCACAGAACCA	871
OY	435	TGCCAGGCTCAAGATTATACAGACACTCAGACAAAGATTAAGAAATCCACACTATGACCA	494
Db	872	CAAGGTTACAGTACTATATGATGATGTGTGTAACACATTAAGTGGAGGCTATGAGAG	931
OY	495	AGCATTTCCGGAATAGGTGAGAGATTAATCTGATTTCCACCATATGACTCTTGTAATT	554
Db	932	GCACCTCAATTAACGGGTTAATGATGAGATTAATGAAATGTGGTATCCAGTTGTAATT	991
OY	555	TTATCTTTCAATTTGAAAACTCAATTCACAAAGATTACATCACAGAAAAGCTTACAAATCG	614
Db	992	CTACCTTCTCTTTGAGAACCTCCATTCACAGAGACTATTTTCATGGAAGAGCTGTTCATATC	1051
OY	615	ATTTTGGCTGGTTCAGTACACTGTTGTCTGAGGTCATCTACGAGAAACCTGTGGAATTA	674
Db	1052	TTTACGGCTGTACGGTTCGTGTGTTACTTGGCCCTCCAGATATACATATGAGGAGTT	1111
OY	675	TATTCACAGTATCATCATTCATCATGTGAGAGATTTAATCTCCACAGTGGTTACCAAA	734
Db	1112	TATACCAAGAACGCCCTTCTCATCATGATGATGATTTCCACATCCCAAGAGAACTGGGGA	1172
OY	735	ATATCTGAAGAACTGACAAAACATATAGTTGTACTTACTTAACTGTGAGAAA	794
Db	1172	TCACCTTAATCTTGGACCAAAAGAGAAACATACAGACAGTATTTTCAACTGAGAAA	1231
OY	795	GGATTTACTGTAAACTCACAGGTTTTTGGAGATCAATGATGCTCGGATGCGACCA	854
Db	1232	ACATTTTATTTTCATGATGAGTTCAATCTTTGGGCTTGAACATGCTGTAGCACTGTGATTA	1291
OY	855	TGTAATAAGGATCAAGATATPAGCTGTGTGAATTTAGAGAAATGGTTTGAATTA	914
Db	1292	TCTTAAAGGAATAGACATATPAGCTATTTAAAGATCTGAAAGGCTGTTCTGGGTTA	1351
OY	915	A 915	
Db	1352	A 1352	

RESULT 12	
AB023628	
LOCUS	1832 bp DNA linear VPR 02-APR-1999
DEFINITION	Danio rerio zfr2 gene for alpha(1,3)fucosyltransferase, complete cds.
ACCESSION	AB023628
VERSION	AB023628.1 GI:4587297
KEYWORDS	alpha(1,3)fucosyltransferase.
SOURCE	Danio rerio DNA.

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

Cypriniformes; Cyprinidae; Danio.
1 (sites)
Kageyama, N., Natsuka, S. and Hase, S.
Molecular cloning and characterization of two zebrafish
alpha(1.3)fucosyltransferase genes developmentally regulated in
embryogenesis
J. Biochem. (1999) In press
2 (bases 1 to 1832)
Kageyama, N., Natsuka, S. and Hase, S.
Direct Submission
Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate
School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034,
Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, tel: 81-6-6850-5381,
fax: 81-6-6850-5383)
Location/Qualifiers
1. 1832
/organism="Danio rerio"
/db_xref="taxon:7955"
332..1510
/gene="zft2"
332..1510
/gene="zft2"
/codon_start=1
/product="alpha(1.3)fucosyltransferase"
/protein_id="BA6767.1"
/translation="MDKLTVMKIKLTPASKSTOKLIVTFMLISVLCIFYVYNNPT
TFPKYPTMGIIHANSSECTELVHKMONTCEIMKNASYALPTAPNPAGKODIIVL
IMAPFPRVFDKDCGLKQVNHGCOLTETRSIIOKAGVMPHHRMSTDPQPRPD
OKVIMNNESTPHCNONDPLNDRPNVTSYRNDSDIPYVGLVDTEOKKFTTPK
DKLVCMIVSNEQHHKRTHEFEPAKHIVSYGGFENNRINDGDIVANVSSCKEFLS
FENSIHRDFTLEKLEPLALGTVPVIGPRVNYEEFIPROAFIHVDLPTPRELADH
LISIDONEBOYKOPFNWRKLYVSKSTFGLHSRCLDYLRKNKHIAVTDLKGMFWG"

BASE COUNT 585 a 394 c 317 g 536 t
ORIGIN

Query Match 24.9%; Score 228.2; DB 5; Length 1832;
Best Local Similarity 55.0%; Pred. No. 5, 6e-50;
Matches 496; Conservative 0; Mismatches 393; Indels 12; Gaps 2;

18 TAACGAACACACATCTGTTGGTGGATGCGCATTTGGGCAACCTTTACCTTACATC 77
619 TCAGACATCATCTGTTGCTGCTGATGCGCTCTTCAAGTACCTTTGACCTGAAAGA 678
78 CTGCCAAGCATGCTCAATATCCAAAGGTCGATCTCACAAAGACCGCTCATTTGACAA 137
679 TTGTGTTGGAGTATTAACATCCACGGCTGTCAATTAACATACGACGAAGTGTGATCCA 738
138 CAATGCCATGCGTCCGATTCACCATAGACATCAGCTGGGATCTGACTAATCTTAC 197
739 GAAGGCGATGAGTATATGTTTACACACAGATATGACACAGAT-----TTTCC 789
198 TCAGCAGGCGACGACCTTTGAGAAATGATTTGATGATTTAGATCACCACCTCA 257
790 GCAACCCCGCGCGCATTTCCAAAAGTGATATGATGATGATGATGATGATGATGATGAT 849
258 CACCCCAAAAAGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317
850 TTGCATATCAAAATGATCTCTTAATATGATGATGATGATGATGATGATGATGATGATGAT 909
318 TTGCATATCAAGTACCTTATGCTTCTTGACGCTGACGACCA--AATCCCTTTGTGT 374
910 TTGCATATCAAGTACCTTATGCTTCTTGACGCTGACGACCA--AATCCCTTTGTGT 969
375 TGAAGTCCCAAGCAGAGAAAGT 434
970 CACCATCCCAAAAAGATTAATAGTTGCTGATGATGATGATGATGATGATGATGATGATGAT 1029
435 TGCAGAGTCAAGTATTTACACAGAGCTCAAGAGATGATGATGATGATGATGATGATGATGAT 494
1030 CAAGCGCACATTTCTTCTACAGATTTGCGAACAACATTAATGATGACCTATGTGAGG 1089

495 AGCATTCGGGAATFACGTACAGATATAAATCTGATCCACCATATCTACTGTAAAT 554
1090 GCACCTTCACAAACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1149
555 TTATCTTCATTTGAAACATCAATTCACAAAGATTAACATCAGAAAGAGCTTCACATCC 614
1150 CTACCTGTCTTTGAGAACCTCAGATTCACAGACTATTTACTGAGAAAGCTGTTCATATCC 1209
615 ATTTTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
1210 TTATGACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
675 TATTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
1270 CATACCAAGAGAGAGCTTCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
735 ATATTCGAAGAGAGTTCACAAACATTAATGATGATGATGATGATGATGATGATGATGATGAT 794
1330 TCACCTTAATTCCTTGACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
795 GGAATTTTACTGTAACCTTACACGCTTTTGGGATGATGATGATGATGATGATGATGATGAT 854
1390 ACTTTATGTTTCAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
855 TGTAAAAGGAGATCAAGATATTAAGTCTGTTGATTAATTAAGAAATGTTTGAATTA 914
1450 TCTTAAAAGGAGATCAAGATATTAAGTCTGTTGATTAATTAAGAAATGTTTGAATTA 1509
915 A 915
1510 A 1510

RESULT 13
AF288369
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

1264 bp DNA linear VRT 02-DEC-2001
Gallus gallus alpha (1.3) fucosyltransferase gene, complete cds.
AF288369.1 GI:17224489

ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1264)
Foley, A.C., Schmitz, B., Stern, C.D. and Streitt, A.
Unpublished
2 (bases 1 to 1264)
Foley, A.C., Schmitz, B., Stern, C.D. and Streitt, A.
Direct Submission
Submitted (19-JUL-2000) Genetics and Development, Columbia
University, 701 West 168th Street, New York, NY 10032, USA
Location/Qualifiers
1. 1264
/organism="Gallus gallus"
/db_xref="taxon:9031"
1..1062
/note="fuc"
/codon_start=1
/product="alpha (1.3) fucosyltransferase"
/protein_id="AAL36994.1"
/db_xref="GI:17224490"
/translation="SMKLTLPILSPFISFLASLNOFRTPEDYNSYISLRLP
AGSTMLPSESMILTLLMTWFGHRENSNSELDPDPCHTFTNRWSQAAADAVIM
HHRDYWDRELSRIPRLPSQWIFNNEPSHSNGLAMDBLFNLTWSYRSDIFP
PYGEOLISEPQPSIPKTKVAVVSNMRRSSHRVYELKHIIPVDYGRDILP
LSMDLHGTQSYNYFLAPENSEHNDYTEKLMRAALSGLTPVYGLPPEVYEHFLP
PDSFTIHVDPSAGELAOYLQSLSDTEKYQKRPQMRWMLKPVILGIANMAYCRACHF
LQRTASVHYVNLKSEVY"

BASE COUNT 286 a 355 c 319 g 304 t
ORIGIN

Query Match 24.1%; Score 220.6; DB 5; Length 1264;
 Best Local Similarity 57.2%; Pred. No. 5.8e-48;
 Matches 441; Conservative 0; Mismatches 324; Indels 6; Gaps 2;

28 ACCATTCCTGTTGGTATGCGCATTTGGGACAGCTTTGACCTTACCTCGGACGA 87
 175 ACCATTCCTGTTGGTATGCGCATTTGGGACAGCTTTGACCTTACCTCGGACGA 234
 88 ATGTCATATCCAAAGGTCGATCCACACAGAGCGCTGATTTGATGATGATGATGAT 147
 235 CTCCTTGACACCCGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 294
 148 GGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
 295 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
 208 AGGCAACCCCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
 355 AGGCTCCCATCCAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 414
 268 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
 415 TTAGTGCATGACGACACCTTCAACCTGATGATGATGATGATGATGATGATGATGAT 474
 328 CAAGTGCCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
 475 TTACACCCCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
 388 AAGGAAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
 535 AAGACCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
 448 TATTCACAGAGCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
 595 TACTACGAGAGCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
 508 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
 652 CCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
 568 GAAACCTCAATTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 624
 712 GAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
 625 GCTTCAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
 772 GGCACCATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
 685 GATTCATTCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 744
 832 GACTCCTTCATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
 745 GAAGTTGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
 892 GAACGACGAGAGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 942

RESULT 14
 AB035907 1092 bp DNA linear VRT 05-JAN-2002
 LOCUS Xenopus laevis gene for xft1, complete cds.
 DEFINITION
 AB035907.1 GI:18146867
 KEYWORDS
 SOURCE
 ORGANISM
 Xenopus laevis
 Xenopus laevis bone marrow DNA.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE
 AUTHORS Kaneo, M., Nishihara, S., Kilano, T., Narimatsu, H. and Saitou, N.
 TITLE The evolutionary history of glycosyltransferase genes

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1092)
 AUTHORS Kaneo, M., Saitou, N. and Kilano, T.
 TITLE Direct Submission
 JOURNAL Submitted (17-DEC-1999) Mika Kaneo, National Institute of
 Genetics, Laboratory of Evolutionary Genetics, Yata 111, Mishima,
 Shizuoka 411-8540, Japan (E-mail: mikane@med.id.yamagata-u.ac.jp,
 Tel: 81-559-81-6790, Fax: 81-559-81-6789)

FEATURES
 source
 1..1092
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /issue_type="bone marrow"
 1..1092
 /gene="xft1"
 1..1092
 /gene="xft1"
 /codon_start=1
 /product="xft1"
 /protein_id="BAB82490.1"
 /db_xref="GI:18146868"
 /translation="MRGSLISVERLRRCRLCRSRRLRSRLAASITCTLVVYV
 CSODLADIPNPLMGTTPPKQVTVIWEPRKRRRLGCDHMLPVYSGHTTNR
 SYQENADIPNPLMGTTPPKQVTVIWEPRKRRRLGCDHMLPVYSGHTTNR
 SYQENADIPNPLMGTTPPKQVTVIWEPRKRRRLGCDHMLPVYSGHTTNR
 DYGRRLDKEDNIVKYSEKRYLAFENSLHDTITKLMNARVATLVGFS
 RINTEMFIRSFTHYDSSPRKRLAFENSLHDTITKLMNARVATLVGFS
 EHTCTQSVKAGNQHRTASDLAWFES"

BASE COUNT 256 a 294 g 261 t
 ORIGIN

Query Match 22.6%; Score 206.6; DB 5; Length 1092;
 Best Local Similarity 56.1%; Pred. No. 3.1e-44;
 Matches 474; Conservative 0; Mismatches 359; Indels 12; Gaps 4;

19 AACGAATCATTCTGTTGGTATGCGCATTTGGGACAGCTTTGACCTTACATCC 76
 199 AAGCAAGTACATGCTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 258
 79 TGCCAGCAAGTTCATATCCAGAGGTCGATGATGATGATGATGATGATGATGATGATGAT 138
 259 TGCCAGTCTCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318
 139 AATTCATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198
 319 GAGGACAGCGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
 199 CAGCAGGCGGCGCCTTCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 258
 376 TTCCGAAAGCGCGCAGAGTCCCAAAAGTATGATGATGATGATGATGATGATGATGATGAT 435
 259 ACCGCCCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318
 436 TGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 495
 319 TCGATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 378
 496 TCGATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
 379 GTGCAAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
 553 CTCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 612
 439 AGGTCAGATATTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 498
 613 AGGTCAGATATTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
 499 TTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 558
 670 TATGCGCTGATATTCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
 559 CTTTCATTTGAATCTCAATTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 615

Db 730 TTGGCTTTGAGATTCTTGGCACGAGATTACATCGAAGAAAGCTGTGGAGAACGCC 789
QY 616 TTTTGGCTGTTGACGACCTGTTGCTGCTGATAGAGAAACATATGAAATTAAT 675
Db 790 TTTAAATCCAAAGCCATCCCATGTCATGGGGCCGACGGCATATACAGAGATGTC 849
QY 676 ATTCCAGCTATTCATTCATGAGAAAGATTTTAAGTCTCCAGTGAATGCAAAA 735
Db 850 ATACCGGCGACGCTCTTCATTCATGAGAGATTTCGCCAGCCAGAGAGCTGGCCATG 909
QY 736 TATCTGAGAGAGATTGCAAAACAAATAGTTACCTTGTACTTAACTGAGAGAA 795
Db 910 TACCTGAGAGACCTGGTTAAACATTCACCTGACGCGAGATTTACAGTGAAGAG 969
QY 796 GATTTTACTGTAACCTACACCGGTTTGGAAATCAATCATGCTGCGATCGACCAT 855
Db 970 CGTTAGACAGCTGACGACGCTCTTCTGAGAGAGATTACGACCGGCTGCGAGAGC 1029
QY 856 GTAAA 860
Db 1030 GTCAA 1034

RESULT 15
GGU73678 2657 bp DNA linear VRT 20-DEC-1996
LOCUS Gallus gallus alpha-1,3-fucosyltransferase (CFT1) gene, complete
DEFINITION cds.
ACCESSION U73678
VERSION U73678.1 GI:1657998
KEYWORDS
SOURCE chicken.
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2657)
AUTHORS Lee, K.P., Carlson, L.M., Woodcock, J.B., Ramachandra, N.,
Schultz, T.L., Davis, T.A., Lowe, J.B., Thompson, C.B., and Larsen, R.D.
TITLE Molecular cloning and characterization of CFT1, a developmentally
regulated avian alpha(1,3)-fucosyltransferase gene
JOURNAL J Biol. Chem. 271 (51), 32960-32967 (1996)
MEDLINE 97115837
REFERENCE 2 (bases 1 to 2657)
AUTHORS Lee, K.P., Carlson, L.M., and Larsen, R.D.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1996) Immune Cell Biology Program, Naval Medical
Research Institute, 8901 Wisconsin Ave., Bethesda, MD 20889, USA

FEATURES
source
1.2657
/organism="Gallus gallus"
/db_xref="taxon:9031"
170..1240
/gene="CFT1"
170..1240
/gene="CFT1"
/note="similar to mammalian fucosyltransferase FucTIV"
/codon_start=-1
/product="alpha-1,3-fucosyltransferase"
/protein_id="AAC60060.1"
/db_xref="GI:1657998"
/translation="MELGRWSPAPRPGCPRRMRRLWALLGLALGNALALVYCVREL
RRGSAAGREGEVTVLLMWEPCRPWRPACRRRYNITCGLSADGRYGEARAVLFH
HRDLALHGRGLPRGPPRPORRWYMMWESPSGLRGLAGLENTMSYRSDVDY
FVPGYIVPPSPRPREVLPKRSIVAMVYNNNEEARVRYROLKEHPIIVYGRG
MALBESVYKITYSAIKFYLATENSQHTDIITKLMKNAPAAASVFPVLDPRRANIERF
IPADSFIVHDDFSPRLATVYLFKFLDKNPSYRFRYFAWRNKYEIVHTSPWDEHYCKVC
EAVRTAGNQLKTVONLAGWFS"

BASE COUNT 581 a 712 c 818 g 546 t
ORIGIN

Query Match 19.1%; Score 175.2; DB 5; Length 2657;
Best Local Similarity 52.9%; Pred. No. 7.1e-36;

Matches 453; Conservative 0; Mismatches 388; Indels 15; Gaps 3;
QY 21 CGAAATACCATTCCTGTTGGTATGCGCATTTGGGACAGCTTTGACCTTACATCTG 80
Db 334 CGAGGTACGAGTGTCTGTGTGTGTGGAGCCCTTGCGGCCCTTGCGGCCGAGCTG 393
QY 81 CCAAGCAATGTTCAATATTCACAGAGGTGCGATCTCACAAACAGCCCTATGTACAA 140
Db 394 CCGAGAGCGCTACAAATACATACAGGGCTGCTCTGAGAGCGGACCGCGGGGATAC 453
QY 141 ATCCATGCGGCTGCTGATTTACCATAGACAT-----CAGCTGGATCTGACTAA 191
Db 454 GGTCTGGGCGCTCTTCCACACCGGACCTGGCGCTGACAGCGCGCCAGGGGCTGCC 513
QY 192 CTTACCTCAGACAGGCCAGGCCACCTTTCAGAAATGATTTGATGATAGATGAC 251
Db 514 CCGCGGGCCCCCGCGGAGACCCCGCGGCGAGCGGTGCTGTGATGATGACTTCAGAT 573
QY 252 CACTCACACCCCGCAAAAGAGTGCAATGAACTTGTTCACCTGACTTAATATG 311
Db 574 CTGCACTCCCGCGGCTGCGGGGCTGCGGGGCTTCTCACTGACATGTCTGAC 633
QY 312 CCGTATTCAGATATCCAAATGCTTATGCTTCTTACGCTGAGCACAATCCCTTTGT 371
Db 634 AGGAGACTCGAGCTCTTCTGCTTACGAGGCTTACGAGCGCGCGCGCGCG 693
QY 372 GTTTGAAGTCCCAAGAGAGAGAGTGTGCTGCTGAGTGTGAGTACTGAAACCTG 431
Db 694 CTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
QY 432 GCATGCCAGGAGTCAAGTATTACAGAGCTCAGCAGAGATTTGAAATCCACACAT 491
Db 754 GCAGCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 813
QY 492 CCAAGCTTGGGAGATAGCTGAGCAATTAATTCATTCCTGCTGCTGCTGCTGCT 551
Db 814 ---GGCGCGGGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
QY 552 ATTTTATCTTCAATTTGAAACATCAATTCACAAAGATTTACATGACGAAAGCT 608
Db 871 GTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
QY 609 CAATGCAATTTTGGTGGTTCAGTACCTGTTGTCTGCTGCTGCTGCTGCTGCTG 668
Db 931 GAAGCGCTTGGCGGACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
QY 669 GAATTAATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 728
Db 991 GCGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
QY 729 AGCAAAATATCTGAGAGATGACAAACAAATTAAGTTGACTTACTTAACTG 788
Db 1051 GGCACCTTACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
QY 789 GAGAAAGATTTTACGTTAAACCTTACACGAGTTTGGGATTCATGATCCTGCTG 848
Db 1111 GCGGAACAAGTATGAGTTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1170
QY 849 CGACACTGTAAAAAG 864
Db 1171 CGAGGCGGTGAGGAGC 1186

Search completed: October 6, 2002, 04:12:59
Job time : 1338.59 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 102.662 Seconds
(without alignments)
15299.444 Million cell updates/sec

Title: US-09-744-748-3_COPY_280_1194

Perfect score: 915
Sequence: 1 acaaaactgattatttaa.....agaatggttggaataa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq.032802:*

1: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	915	100.0	2036	21	AAZ92672
2	915	100.0	2170	21	AAZ92645
3	791.8	86.5	2676	21	AAZ92647
4	791.8	86.5	2822	21	AAZ92646
5	129	14.1	1814	18	AAT59505
6	128.6	14.1	3594	18	AAT85219
7	128.2	14.0	1086	15	AAO56912
8	128.2	14.0	2042	18	AAT80111
9	128.2	14.0	2043	12	AAQ13330

10	128.2	14.0	2043	15	AAQ56906
11	128.2	14.0	2043	18	AAT76765
12	128.2	14.0	2043	18	AAT61675
13	126.6	13.8	1316	15	AAO56911
14	126.6	13.8	1316	18	AAT61679
15	124.6	13.6	1654	15	AAO56905
16	124.6	13.6	1654	18	AAT61680
17	123	13.4	1126	21	AAZ21135
18	123	13.4	1126	21	AAZ21135
19	123	13.4	6944	21	AAZ21137
20	123	13.4	6944	21	AAZ21137
21	122.8	13.4	1701	15	AAQ77732
22	122.8	13.4	1701	15	AAZ21136
23	122.8	13.4	1701	21	AAZ35014
24	120	13.1	1256	21	AAZ21134
25	120	13.1	1256	21	AAZ35012
26	120	13.1	1400	13	AAO31436
27	120	13.1	1488	15	AAO56910
28	120	13.1	2175	12	AAO14382
29	120	13.1	2175	12	AAZ17082
30	120	13.1	2861	12	AAO14383
31	120	13.1	2861	12	AAZ21133
32	120	13.1	2861	21	AAZ35011
33	120	13.1	2861	21	AAZ17083
34	120	13.1	3647	12	AAO13333
35	120	13.1	3647	15	AAO56909
36	120	13.1	3647	18	AAZ61678
37	118.4	12.9	2175	11	AAO06691
38	116.8	12.8	2134	18	AAT59506
39	116	12.7	795	17	AAT13798
40	64.6	7.1	1578	23	ABL11867
41	64.6	7.1	4229	23	ABL11866
42	53	5.8	1353	21	AAZ51685
43	43.6	4.8	2198	21	AAZ97937
44	40.8	4.5	398	22	AAZ64428
45	40.8	4.5	676	22	AAZ41344

ALIGNMENTS

RESULT 1
AAZ92672
ID AAZ92672 standard; cDNA: 2036 BP.
XX
AC AAZ92672;
XX
DT 05-JUN-2000 (first entry)
XX
AC
XX
DE Murine alpha-1,3-fucosyltransferase gene exon 1 DNA, SEQ ID NO:30.
XX
XX
KW Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW brain; kidney; recombinant expression; transgenic animal; knockout
KW animal; FDC-TIV; drug screening; inhibitor; potentiators; diagnosis;
KW treatment; cancer; murine; mouse; ss.
XX
OS
XX
PN Mus sp.
XX
PN WO200006708-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-JP04092.
XX
PR 29-JUL-1998; 98JP-0213823.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Narimatsu H, Kudo T, Sasaki K;
XX
DR WPI, 2000-183120/16.
XX
PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x

PS Claim 4; Page 127-134; 172pp; Japanese.

XX The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA292645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA292645), and sequences AA292670-292673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).

XX Sequence 2170 BP; 715 A; 395 C; 368 G; 692 T; 0 other;

Query Match 100.0%; Score 915; DB 21; Length 2170;
 Best Local Similarity 100.0%; Pred. No. 9,7e-261;
 Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAAAGCTGATTTATTTTAAAGAACTACCATTCGTTGGTATGCGCATTTGGGCGAG 60
 DB 280 ACAAAGCTGATTTATTTTAAAGAACTACCATTCGTTGGTATGCGCATTTGGGCGAG 339
 QY 61 ACCTTGGACCTTACATTCCTGCCAAGCAATTTTCAATTCGAAGGCGCATTCACACA 120
 DB 340 ACCTTGGACCTTACATTCCTGCCAAGCAATTTTCAATTCGAAGGCGCATTCACACA 399
 QY 121 GACCGCTCATTTGTAACAACAATCCATGCGCTCTGATTCACCATAGAGCATCAGCTGG 180
 DB 400 GACCGCTCATTTGTAACAACAATCCATGCGCTCTGATTCACCATAGAGCATCAGCTGG 459
 QY 181 GATCTGACTACCTTACCTGAGAGGCCAGGCCACCTTTGAGAAATGGATTTGGATGAAT 240
 DB 460 GATCTGACTACCTTACCTGAGAGGCCAGGCCACCTTTGAGAAATGGATTTGGATGAAT 519
 QY 241 TTAGAGTCACCCACCTACACACCCCAAGAGAGTGCGATTAAACATTTGTTCAACCTGACT 300
 DB 520 TTAGAGTCACCCACCTACACACCCCAAGAGAGTGCGATTAAACATTTGTTCAACCTGACT 579
 QY 301 CTAACTTATGCGCGTATTCAGATTCACAAGTCCCTTATGGCTTTCGACGGTGAGCACA 360
 DB 580 CTAACTTATGCGCGTATTCAGATTCACAAGTCCCTTATGGCTTTCGACGGTGAGCACA 639
 QY 361 AATCCCTTTGTGTTGAAGTGCCAAAGCAGAGAGAGTGGTGTGCTGGGTTGAGTAAAC 420
 DB 640 AATCCCTTTGTGTTGAAGTGCCAAAGCAGAGAGAGTGGTGTGCTGGGTTGAGTAAAC 699
 QY 421 TGAACCCCTGAGAGTCCAGAGGTCAAGTATTTACAAGAGCTGACGAAGAGTATTGAATC 480
 DB 700 TGAACCCCTGAGAGTCCAGAGGTCAAGTATTTACAAGAGCTGACGAAGAGTATTGAATC 759
 QY 481 CACACCTATGCGCAGCATTCGAGAAATACGTAAACGATTTAAATTCGATTCCACACATA 540
 DB 760 CACACCTATGCGCAGCATTCGAGAAATACGTAAACGATTTAAATTCGATTCCACACATA 819

QY 541 TCTACTGTAAATTTATATCTTTTCATTTGAAAACCTCAATTCACAAGATTACATCAGACAA 600
 DB 820 TCTACTGTAAATTTATATCTTTTCATTTGAAAACCTCAATTCACAAGATTACATCAGACAA 879
 QY 601 AAGCTTACATATGATTTTGGCTGGTACGTACCTGTTTCTGGGTCATATAGGAA 660
 DB 880 AAGCTTACATATGATTTTGGCTGGTACGTACCTGTTTCTGGGTCATATAGGAA 939
 QY 661 AACTATGAGATATATATTCAGCTGATTCATTCATTCATGAGAGATTTTAACCTGCC 720
 DB 940 AACTATGAGATATATATTCAGCTGATTCATTCATTCATGAGAGATTTTAACCTGCC 999
 QY 721 AGTAGATTACCAAAATATCTGAAGAGTTGACAAAAACAATAGTTGACTTACTTAC 780
 DB 1000 AGTAGATTACCAAAATATCTGAAGAGTTGACAAAAACAATAGTTGACTTACTTAC 1059
 QY 781 TTATACCTGGAGAAAGATTTTACTGTAACCTTACACCGTTTGGGATTCACATGATGC 840
 DB 1060 TTATACCTGGAGAAAGATTTTACTGTAACCTTACACCGTTTGGGATTCACATGATGC 1119
 QY 841 CTGCGATGCGACCATGTAATAAGCATCAAGATTTAAGTCTGTGTAATTAGAGAA 900
 DB 1120 CTGCGATGCGACCATGTAATAAGCATCAAGATTTAAGTCTGTGTAATTAGAGAA 1179
 QY 901 TGGTTTTTGGAAATTA 915
 DB 1180 TGGTTTTTGGAAATTA 1194

RESULT 3
 AA292647
 ID AA292647 standard; cDNA; 2676 BP.

XX AA292647;
 AC
 XX
 DT 05-JUN-2000 (first entry)

XX Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.

XX Alpha-1,3-fucosyltransferase: fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentiators; diagnosis;
 KW treatment; cancer; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS /tag= a
 FT /product= "Human alpha-1,3-fucosyltransferase"

PN MO200006708-A1.

PD 10-FEB-2000.

PF 29-JUL-1999; 99WO-JP04092.

PR 29-JUL-1998; 98JP-0213823.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Narimatsu H, Kudo T, Sasaki K;

DR WPI: 2000-183120/16.

DR P-PSDB: AY880936.

PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT epitope and an antibody recognizing it useful for diagnosis of brain
 PT and kidney disease and cancer.

PS Claim 4; Page 143-150; 172pp; Japanese.

XX The invention relates to a novel alpha-1,3-fucosyltransferase which

transfers a fucose moiety to galactosyl-beta1-4-N-acetylglucosamine (generating the Lewis x or y epitope). It does not transfer a fucose moiety to alpha-2/3-sialyl-galactosyl-beta1-4-N-acetylglucosamine and therefore does not generate the sialyl-Lewis x epitope. The invention also relates to DNA sequences encoding alpha-1,3-fucosyltransferase and expression vectors and host cells comprising these DNA sequences. The invention additionally encompasses the preparation of alpha-1,3-fucosyltransferase via the culture of transformed cells or by expression of the protein in a transgenic animal; antibodies which recognise alpha-1,3-fucosyltransferase; methods for screening potential inhibitors or potentiators of alpha-1,3-fucosyltransferase activity or expression; the preparation of compounds having fucose-containing sugar chains by use of the protein; and knockout non-human animals lacking alpha-1,3-fucosyltransferase. Alpha-1,3-fucosyltransferase has a broad substrate range to the known FUC-RTV and is expressed mainly in brain and kidney tissues. Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies, potentiators and inhibitors may be used for the treatment and diagnosis of diseases of the brain and kidney, and/or of cancers. They may be used for the identification of substances which affect the activity or expression of alpha-1,3-fucosyltransferase, such substances may be used therapeutically. The knockout animals can be used to study the mechanisms of action and expression of alpha-1,3-fucosyltransferase. Sequences AA632646 and AA632647 represent cDNAs encoding human alpha-1,3-fucosyltransferase (AA6780996).

SQ Sequence 2676 BP; 819 A; 527 C; 476 G; 854 T; 0 other.

Query Match	86.58;	Score 791.8;	DB 21;	Length 2676;
PostgreSQL	01.68;	Score 791.8;	DB 21;	Length 2676;

Matches	838;	Conservative	0;	Mismatches	77;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	-----	--------	----	------	----

OY	1	ACAAAAGTATTTTATTTTAAAGAAATACACATTCGCGTGGGTATGGCCATTTGGGCAG	60
Db	460	ACCAAACTGATTTTATTAATGAACATACATATTCGTGGTGGTGTGGCCATTTGGGCAG	519
OY	61	ACCTTTGACCTTACATCCGCGAAGAAATGTTCAATATTCGAAGAGGTGGCATCTCACACA	120
Db	520	ACCTTTGACCTTACATCTTCGCGAAGAAATGTTCAATATTCGAAGAGGTGGCATCTCACAG	579
OY	121	GACCGCTCATGTTACAAACAATCCCATGCGGTCCTGATTACCATAGACATCAGCTGG	180
Db	580	GACCGTTACGTACAAACAATCCCATGCGGTTCTATTCATCACCGACCATCAGTTGG	639
OY	181	GATCTGACTAATCTACCTGACGAGGCCACCCCTTTCAGAAATGATTTGGATCAAT	240
Db	640	GATCTGACAAATTTACTCTGAGCAAGTAGGCCACCTTCAGAAATGATTTGGATCAAT	699
OY	241	TTAGAGTACACCACATCTCACACCCCAAGAGGCGATTGAAACATTTTCAACCTGACT	300
Db	700	TTGGAATACCAATCTCACACTCCCCAAAGAGTGGGATTGAGCAATTTGTTAACTGACT	759
OY	301	CTAACTTATGCGCGGTATTCAGATATTCGAAGGCGCTTATAGGCTCTTAGCGGTACAGACA	360
Db	760	CTGACTTATACGCCCTGATTTAGATATTCGAATGCCATTCATAGGCTTCTTAGCGGTACAGACA	819
OY	361	AATCCCTTTGTGTTGAAGTCCACAGACAGAAAGTTGTGTGTCTGGGTTGTGATGATAC	420
Db	820	AATCCCTTCGTGTTGAAGTCCACAGACAGAAATGTGTGTGTGGTGTGTGATGATAC	879
OY	421	TGGAAACCGTGCATGCGAGGTCAGATATTAACAACGAGCCTGACCAAGATTTGAATAC	480
Db	880	TGGAAACCGTGCATGCGAGATTCAGATATTAACAATGAGCTTACGAAAGAGATTGAATAC	939
OY	481	CACACATATGGCCAAAGCATTCGAGAAATACGGAACGATATTAATAATCTGATCCCAACATA	540
Db	940	CATACCTACGGCCAGACATTTGAGAAATATGTCATATGATATTAATAATCTGATCCCAACATA	999
OY	541	TCCTACTGTAAATTTATCTTTCAATTTGAAAATCTCAATTCGCAAAAGATTTACATACAGAA	600
Db	1000	TCCTGCTGTAAATTTATCTTTCTCCTTTGAAAATTCATATCCACAGAGATTTACATACAGAA	1059
OY	601	AAGCTTCACAAATGCATTTTGGCGTGTAGTACCTGTTTCTCTGGGTCCATCTTAGGAA	660

Db	1060	AACCTTACAAACGCTTTTCCTGGCTCGTACTGTTGTTCGGAGCATTGAGGAA	11139
Qy	661	AACATATGAAATTAATTTCCAGCTGATTAATCATTCATGTGGAGAAATTTAACTCTCC	720
Db	1120	AACATAGAAATTAATTAATTCACGACGATCATTCATTCATGAGGAAGTATTAACCTGCC	1179
Qy	721	AGTAGTATGCAAAATATCTGAGAGAGTTGACAAAAACAATAGTTGTAACTTAGTAC	780
Db	1180	AGTAGACTGCAAAATATCTGAGAGAGTGCACAAAAACAATAGTTATACCTTAATTAC	1239
Qy	781	TTTAACTGAGAAAGATTTTACTGTAACTACACAGCTTTTGGGAATACATGTCATAC	840
Db	1240	TTTAACTGAGAGAAAGATTTTACTGTAACTTCCACAAATTTGGGAATCACATGATCT	1299
Qy	841	CTTGCACTGGACATCTGTAAAAAGGCATCAAAATTAAGTGTGGTAAATTTAGAGAA	900
Db	1300	TTTGCTTGGCAGATCATGTAAAAAGCATTAAATTAAGTGTGGTAAATTTAGAGAA	1359
Qy	901	TGCTTTTGGCAATTA 915	
Db	1360	TGGTTTGGCAATTA 1374	

RESULT 4

ID AAZ92646 standard; cDNA; 2822 BP.

AC AAZ92646;

DT 05-JUN-2000 (first entry)

DE Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:4.

KW Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;

KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;

How do we do it?

XX
XX
XX

FT	1289.2368
CDS	44.1

FT	
/product=	"Human alpha-1,3-fucosyltransferase"

PN WO200006708-A1

PD 10-FEB-2000.

PF 29-JUL-1999;

PR 29-JUL-1998;

PA (KYOW) KYOW
 YV

PI Narimatsu H,
yy

DR WPI; 2000-18

XX
XX
Alpha 1 3-4

PT epitope and

XX
L
g

XX
XX
E

CC transfers a

CC moiety to al

CC invention al

Claim 4; Page 134-142; 172pp; Japanese.

The invention relates to a novel α -1,3-fucosyltransferase which

(generating the Lewis x or y epitope). It does not transfer a fucose

and therefore does not generate the sialyl-Lewis x epitope. The

SQ Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other:
 Query Match 14.1%; Score 129; DB 18; Length 1814;
 Best Local Similarity 52.38; Pred. No. 9e-28;
 Matches 335; Conservative 0; Mismatches 300; Indels 6; Gaps 2;

```

OY 104 GGTGCAATCTCACACAGACCGCTCATTTGACAAACAATCCATCGGCTTATGACC 163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 548 GCTGCGCTGATGCTTACCGGAGCGCTGTACCGAGTGTGCTGCTGCTTACAC 607
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 164 ATAGAGATCATGAGTGGATCTGATTAATTAATTAATTAATTAATTAATTA 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 608 ACCGAGACTCTCAAAACCGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 667
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 224 AATGATTTTGATGATTAATTTAGTCAACCCACTCACACCCCAAAAGATGGA 283
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 668 CTGCGCTGCGCTCTCCATGGAATCGCCAGTAATACCATGCTCTCCATGCG 727
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 284 ACTTGTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 343
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 728 GCATCTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 787
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 344 TCTTGAACGCTGACGCAAAATCCCTTTGCTTTGAATGCTGCAAGCAAGAA 403
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 788 GCTTGGAGCTCTCTCTGCGGCC---ACATCCCACTACCGGCAAAAGCAAG 844
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 404 GCTGGGTTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 463
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 845 CCTGGGTTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 904
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 464 GCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 523
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 905 CCCCTCATCTGAGTGTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 524 ATCTGATTTCCACCATTAATTAATTAATTAATTAATTAATTAATTAATTA 583
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 965 GTCTGCTGCCACATTTGGCCGGTACCGGCTTACTGCTGCTTGAAGTACAG 1024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 584 AAGATTACATCAGCAAAAGCTCTA---CAATGCAATTTTGGCTGCTGCTG 640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1025 GGGACTACATCAGTGAAGATGCTGGCGCAATCCCTGGCGCTGCTGCTG 1084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 641 TCTTGGTTCATCTAGGAAATCTAGGAATTAATTAATTAATTAATTAATTA 700
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1085 GCGTGGGACCTCTCGGCCCACTACGAGGCTTTTGTGCCACCATGATCC 1144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 701 TGAAGATTTTAACTCTCCAGTGAATTAATTAATTAATTAATTAATTA 741
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1145 TGAAGACTTCACTGCTGCGGTAAGTGTGCTGCTGCTGCTGCTGCTG 1185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```

RESULT 6
 AAT85219 standard; DNA: 3594 BP.
 XX AAT85219;
 AC
 XX
 DT 02-MAR-1998 (first entry)
 XX
 DE Mouse alpha-fucosyltransferase Fuc-TVII gene.
 XX
 KW Alpha-fucosyltransferase; Fuc-TVII gene; mouse;
 KW sialyl Lewis x determinant; oligosaccharide; antiinflammatory;
 KW inhibitor; ds.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 996..3082
 FT /*tag= a
 FT /note= "includes introns"
 FT exon 465..477
 FT /*tag= b

FT /number= 1
 FT 478..669
 FT /*tag= c
 FT /number= 1
 FT 670..1149
 FT exon
 FT /*tag= d
 FT /number= 2
 FT /codon_start= 996..998
 FT 1150..1946
 FT /*tag= e
 FT /number= 2
 FT 1947..1959
 FT /*tag= f
 FT /number= 3
 FT /codon_start= 1947..1949
 FT /note= "alternative ATG start codon"
 FT 1960..2065
 FT /*tag= g
 FT /number= 3
 FT 2067..3552
 FT /*tag= f
 FT /number= 4
 FT /codon_start= 2126..2128
 FT /note= "alternative ATG start codon"

W09732889-A1.
 12-SEP-1997.
 07-MAR-1997: 97WO-US03184.
 08-MAR-1996: 96US-0613098.
 (UNMI) UNIV MICHIGAN.
 Gersten KM, Lowe JB, Natsuka S;
 WPI: 1997-470535/43.
 P-PSDB; AAW27138 and AAW26671.
 Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g.
 screening for inhibitors used as antiinflammatory agents
 Claim 1: Fig 2: 91pp: English.

This DNA sequence comprises the mouse Fuc-TVIII gene encoding
 alpha-fucosyltransferase. It was isolated from a 373 cell genomic
 library by screening with a probe from the catalytic domain of the
 human Lewis alpha(1,3/1,4)fucosyltransferase (Fuc-TIII).
 Examination of the mouse Fuc-TVIII locus identified 3 Met codons
 that may initiate translation of alpha(1,3) fucosyltransferases
 with different cytosolic domains encoded by exons 2 and/or 3, but
 with identical Golgi-localised catalytic domains encoded by exon 4
 (see AAW27138 and AAW26671). The proteins and nucleic acids can
 be used to construct animal cell lines able to post-translationally
 modify oligosaccharides on cell surface, intracellular or secreted
 proteins and lipids, to isolate reagents for efficient enzymatic
 production of oligosaccharides, to generate antibodies to
 glycosyltransferases useful as diagnostic reagents, to screen for
 fucosyltransferase inhibitors and inactivators, especially those
 that act as antiinflammatory agents, for genotyping individuals at
 the fucosyltransferase locus and for in vitro synthesis of sialyl
 Lewis x tetrasaccharide.

SQ Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other:
 Query Match 14.1%; Score 128.6; DB 18; Length 3594;
 Best Local Similarity 52.5%; Pred. No. 1.6e-27;
 Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

```

OY 104 GGTGCAATCTCACACAGACCGCTCATTTGACAAACAATCCATCGGCTTATGACC 163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2277 GCTGCGCTGATGCTTACCGGAGCGCTGTACCGAGTGTGCTGCTTACAC 2336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```


XX P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;
 KW calcium-dependent carbohydrate binding protein; platelet; endothelium;
 KW thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;
 KW inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
 KW disseminated intravascular coagulation; rheumatoid arthritis; antibody;
 KW adult respiratory distress syndrome; tumour metastasis;
 KW fucosyltransferase III; acetylglucosaminyltransferase;
 KW autoimmune disorder; inflammatory disorder; ss.

OS Homo sapiens;
 XX MO9706176-A2.
 XX

XX 20-FEB-1997.
 XX

XX 02-AUG-1996; 96WO-US12820.
 XX

XX 17-MAY-1996; 96US-0649802.
 XX

XX 03-AUG-1995; 95US-0310920.
 XX

XX 13-MAY-1996; 96US-0017794.
 XX

XX (OKLA) UNTV OKLAHOMA STATE.
 XX

XX Cummings RD, McEever RP, Moore KL;
 XX

XX WPI; 1997-154206/14.
 XX

XX New O-glycan(s) derived from P-selectin glycoprotein ligand
 XX (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
 XX for treating inflammatory or autoimmune disorders or tumours

XX Disclosure; Page 80-81; 96pp; English.

XX This sequence represents the coding sequence for the human
 XX fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-
 XX acetylglucosaminyltransferase encoded by AA0180112 are needed for the
 XX expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in
 XX mammalian cells. P-selectin is a calcium-dependent carbohydrate binding
 XX protein expressed on the surfaces of activated platelets and endothelium
 XX in response to thrombin and other agonists. PSGL-1 (see AA026174) is a
 XX high affinity P-selectin ligand produced by leukocytes. Binding of
 XX P-selectin to PSGL-1 is calcium ion dependent and is abolished by
 XX treatment of the ligand with sialidase. PSGL-1 is a homodimer, with a
 XX highly extended extracellular domain, which is a feature of mucin-like
 XX proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of
 XX the invention (see AA026175-026188). PSGL-1 is also targeted by the
 XX O-glycans, and O-glycanated products of the invention. The products are
 XX used for inhibiting the binding of P-selectin and other selectins to
 XX PSGL-1. They can be used for the treatment of inflammation, injury
 XX resulting from ischaemia and reperfusion, bacterial sepsis and
 XX disseminated intravascular coagulation, adult respiratory distress
 XX syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and
 XX other autoimmune or inflammatory disorders. The products can also be used
 XX in the production of antibodies and in diagnostic applications.

XX Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

XX Query Match 14.08; Score 128.2; DB 18; Length 2042;
 XX Best Local Similarity 53.08; Pred. No. 1.6e-27;

XX Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

XX 106 TGGCATCTGCACACAGACCGCTCATGTGTACACAAATCCCATCGCTCTGTATTCACAT 165

XX 343 TGGCATCTGCACACAGACCGCTCATGTGTACACAGCGGTATGTATGTACACAC 402

XX 166 AGAGACATCAGCTGGGA---TCGTACTTAATTTACTCCACAGCGGACACCCCTTTG 222

XX 403 TGGGATATATCATGTCCAAACCCCTAAGTCAAGCTCCACCTTCCCGAGGCGGACAG 462

XX 223 AAATGATTTGATGATTAATTTAGAGTCAACCCACACACCCCAAAAGAGTGACATTTGA 282

XX 463 CGGTGATCTGTTCACCTTGAAGCACCCCTTACTGTCCACACCTTGAGACCTTGAGAC 522

QY 283 CACTTGTCAACCTGACTCTACTTATTCGGCGTGAATTCAGATATTCAGATGCTTATGCG 342
 DB 523 AGATTACTCATCTCAACATGCTCTACCGCAGGACCTCGAGATCTTACGCCCTTGCGG 582
 QY 343 TTTGTGACGGTGAACCAATTCCTTTG-----TGTGAAGTCCCAAGCAAGGAG 393
 DB 583 TGGTGGAGGCGGTGGTGGCGGACCGCCACCCACCGCTTACCTTGGCGGACAGCC 642
 QY 394 AAGTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 453
 DB 643 GAGCTGTGGCTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 702
 QY 454 AAGGAGCTGACGAAAGATTTGAATTCACACCTATGCGCAAGCAATTCGGAATTCG 513
 DB 703 CAGAGCTTGAGGCTCATCTCAAGTGAAGCTGTACGACGCTCCGACA---AGCCCTG 759
 QY 514 AAGCATTAATTCGATTCCTCCACCATATCTACTTGTAAATTTATCTTCTTGAATAC 573
 DB 760 CCGAAGGAGCGCATGATGAGACGCTGTCCGGTACAAAGTTCTACTGCTTGAGAAC 819
 QY 574 TCAATTCACAAAGATTACATCAGAGAAAGCTCT---ACATGCAATTTTGGCTGTGA 630
 DB 820 GTGCCCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 879
 QY 631 GTACCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 690
 DB 880 GTGCCCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 939
 QY 691 TTTCAATATGAGAAATTTTAACTTCTCCAGATGATTTGCAAAATATCTAAGAAAGTT 750
 DB 940 TTTCAATATGAGAAATTTTAACTTCTCCAGATGATTTGCAAAATATCTAAGAAAGTT 750
 QY 751 GACAAAAACAAATTAAGTTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 795
 DB 1000 GACAAAGGACCCAGCCCGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCT 1044

RESULT 9
 ID AA013330 standard; DNA; 2043 BP.
 XX AA013330;

XX 07-NOV-1991 (first entry)

XX Glycosyltransferase gene in clone pCDM7-alpha(1,3/1,4)FT.

XX ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 73..1158

XX polyA_site 2023..2034

XX polyA_signal 1926..2003

XX repeat_region 1183..1547

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

XX repeat_region 1885..1999

PR 14-FEB-1990: 90US-0479858.
 PR 14-FEB-1990: 90US-0480133.
 PA (UNMI) UNIV OF MICHIGAN.
 XX
 XX
 PI Lowe JB;
 XX
 DR WPI: 1991-267151/36.
 DR P-PSDB: AAR13749.
 XX
 PT Isolation of gene conveying post-translational characteristic -
 PT e.g. the presence of soluble or membrane bound oligo or
 PT polysaccharide or glycosyltransferase.
 XX
 PS Claim 11: Fig 1: 155pp: English.
 XX
 CC The DNA encodes a protein sequence capable of functioning as a
 CC GDP-Fuc:[beta-D-Gal(1,4/1,3)]-D-GlcNAc(6/3)alpha (1,3/1,4)-
 CC fucosyltransferase. The sequence coded by nucleotides 127 to 1083
 CC represents the functional protein. The DNA sequence can be used to
 CC construct animal cell lines with respect to posttranslational
 CC modification of the oligosaccharides on cell-surface, intracellular
 CC or secreted proteins or lipids by sub-terminal alpha(1,3) and alpha
 CC (1,4) fucose residues that represent the products of this enzyme.
 CC These can be used for the prodn. of diagnostics and therapeutics.
 CC See also AAQ13331-013333.
 CC
 XX
 SQ Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other;

Query Match 14.0%; Score 128.2; DB 12; Length 2043;
 Best Local Similarity 53.0%; Pred. No. 1.6e-27;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 106 TGGCATCTCAACAACAGCCGCTATTGTACAAACAAATCCCATGCGTCTGATTACCAT 165
 DB 343 TGGCAATCATCTGCGCCAGCGGTGTACCCAGGCAACAGCGCATCGTGACAC 402
 QY 166 AGAGACATCATGCTGGGA---TGTGACTAACTTACCTCAGAGCGCCACCTTTGAG 222
 DB 403 TGGGAATCATCTGCAACCCCTTAAGTCACGCTCCACCTTCCCGAGGCGGAGG 462
 QY 223 AAATGGATTGGATGATTAAGTACACCCACACACCCCAAAAGATGGCATTTGAA 282
 DB 463 CGCTGGATCTGTTGTTCAACTTGGAGCCACCCCTTAAGTCAGAGCGCTGGAG 522
 QY 283 CACTTGTCAACCTGACTTAATTCGCGCTGATTCAGATTCACAAAGCTTATGCG 342
 DB 523 AGATGACTTCAATCTCACCAATGCTCTACCGAGGACTCCGACATCTTCACGCCCT 582
 QY 343 TTTCTGACGGTGAACCAAAATCCCTTTG-----TGTGTAAGTCCCAAGCAAGAG 393
 DB 583 TGGCTGGAGCCGTGTCCGCGCAGCTGCCACCCAGCTCAACCTCTGGGCAAGAGC 642
 QY 394 AATTTGTGTGCTGGTGTGGATGATAACCTGAGATCCGAGGCTCAAGTATTAC 453
 DB 643 GAGCTGTGGCTGTGGGGGTGCTCAACTGGAACCGGACTCAACGAGGCGCTCTAC 702
 QY 454 AAGAGCTCAGCAAGATTAATTAATTCACACCTATGCGCAACCATTCGAGATAGTG 513
 DB 703 CAGAGCTGAGGCTCATCTCAAGGTGAGCGTGTAGGAGCGCTCCACA---AGCCCTG 759
 QY 514 AACGATTAATAATCTGATTCACCATATCTACTTGTAAATTTTATCTTTCAATTTGAA 573
 DB 760 CCCAAGGGGACATGATGAGAGCGTGTCCCGGTACAAAGTTTACCTGGGCTTCGAGAAC 819
 QY 574 TCAATTCACAAAGATTAATTAATTCACAGAAAGCTCT---ACATGCAATTTTGGTGTG 630
 DB 820 TCGTTGACCCCGACTACATACACGAGAGCTGTGAGGAGACCGCTGAGGCGTGGGCC 879
 QY 631 GTACCTGTTGCTCTGGTCCATCTAGGGAATAATATATATTCAGAGCTGATTCA 690
 DB 880 GTGCCCGTGTGTGTGGGCCCCACAGCAAGCAACTAGAGAGGTTCTCTGCCACCGCAAGCC 939

QY 691 TTGATTCATGTGGAAGATTTTAACTCTCCAGTGAAGTACCAAAATATCTGAAGAGATT 750
 DB 940 TTGATTCATGTGGAAGATTTTAACTCTCCAGTGAAGTACCAAAATATCTGAAGAGATT 999
 QY 751 GACAAAAACATAAGTGTGACTTACTTACTTAACTGAGAAAG 795
 DB 1000 GACAAAGACCAACGCCCGCTTACTGAGCTTTCGCTGGCGGAG 1044

RESULT 10
 AAQ56906
 ID AAQ56906 standard; DNA: 2043 BP.
 AC AAQ56906;
 XX
 XX
 DT 26-JUL-1994 (first entry)
 XX
 DE DNA encoding a glycosyltransferase.
 XX
 KW Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
 KW surface; oligosaccharide; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 72..1158
 FT /tag= a
 XX
 PN M09402616-A.
 XX
 PD 03-FEB-1994.
 XX
 PF 20-JUL-1993: 93WO-US06703.
 XX
 PR 20-JUL-1992: 92US-0914281.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Lowe JB;
 XX
 DR WPI: 1994-048874/06.
 DR P-PSDB: AAR45934.
 XX
 PT DNA fragment encoding a glycosyltransferase - can be used for in
 PT vitro reactions to modify cell surface oligosaccharide(s) e.g.
 PT blood gp. determinants, to protect against transplant rejection
 XX
 PS Disclosure: Fig 1: 249pp: English.
 XX
 CC The sequence is that encoding human glycosyl transferase. The enzyme
 CC produced by the DNA may be non glycosylated. This prevents premature
 CC loss of enzyme activity. It can also be used in in vitro reactions to
 CC modify cell surface oligosaccharide mols. e.g. blood group determinants.
 CC See also AAQ56905-12.
 CC
 XX
 SQ Sequence 2043 BP; 413 A; 647 C; 575 G; 408 T; 0 other;

Query Match 14.0%; Score 128.2; DB 15; Length 2043;
 Best Local Similarity 53.0%; Pred. No. 1.6e-27;
 Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 106 TGGCATCTCAACAACAGCCGCTATTGTACAAACAAATCCCATGCGTCTGATTACCAT 165
 DB 343 TGGCAATCATCTGCGCCAGCGGTGTACCCAGGCAACAGCGCATCGTGACAC 402
 QY 166 AGAGACATCATGCTGGGA---TGTGACTAACTTACTCTACAGAGCGCCACCTTTGAG 222
 DB 403 TGGGAATCATCTGCAACCCCTTAAGTCACGCTCCACCTTCCCGAGGCGGAGG 462
 QY 223 AAATGGATTGGATGATTAAGTACACCCACACACCCCAAAAGATGGCATTTGAA 282
 DB 463 CGCTGGATCTGTTGTTCAACTTGGAGCCACCCCTTAAGTCAGAGCGCTGGAG 522


```
QY 283 CACCTGTTCAACCTGACTCTAATCTATGCGCCGATTCAGATATCCAACTGCTTATGCG 342
DB 523 AGATACCTTCATCTCACACATGTCCTACCGACGACGACCTCCACATCTCTTCAAGCCCTACGCG 582
QY 343 TTCTTGACGGTGAGCACAAATCCCTTTG-----TGTTCGAAGTGCACGAAGAG 393
DB 583 TGGCTGGAGCCGTGGTCCGGCCAGCCTGCCACCACCGGCTCAACCTCTGGCGCAAGACC 642
QY 394 AAGTTGGTGTGCTGGTGTGGTGAATCTGGAATCCCTGAGCATGCGAGGGTCAAGATATAC 453
DB 643 GAGCTGGTGGCTGGGGGGGTGTCACAACTGGAAGCGGACATCAGCAGGGGTGGCTATCTAC 702
QY 454 AACGAGCTCAGCAAGAGATGTAATCCACACCTATGAGCCCAACGATTCGGAGAAATACG 513
DB 703 CAGAGCTGCAGGCTCATCTCAAGGTGGAGCTGAGGAGCTGCCACA---AGCCCTG 759
QY 514 AACGATAAAATCTGATTCGCCACATCTACTGTAAATTTATCTTCAATTTGAAAC 573
DB 760 CCCAAGGGGACCATGATGAGAGAGCGCTGCCGTACAAAGTTCTACCTGCGCTTGGAGAAC 819
QY 574 TCAATTCACAAAGATTACATACAGAAAGCTCT---ACAATGCAATTTTGGCTGTTCA 630
DB 820 TCCCTGACACCCCGACTACATCACCGAAGGCTGAGAGAAAGCCCTGGAGGCTGGGCG 879
QY 631 GTACCTGTGTGCTGGTGTGCTGATCTATGAGGAAACCTATGAGAAATATATCCAGCTATTC 690
DB 880 GTGCCCGTGGTGGTGGGCGCCAGCAGAAAGCAACTACGAGAGTTCTGCGCCCGACGCC 939
QY 691 TTCATTCATGTGGAAGATTTTAATCTCTCCAGTGAATGACAAATATCTGAAGAAAT 750
DB 940 TTCACTCCAGCTGAGCAGCTTCAGACCCCAAGGACCTGGCCGAGTCTTGGAGAGCTG 999
QY 751 GACAAAAACAATAAGTGTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 795
DB 1000 GACAAAGACACGCCGCTACCTGAGCTACTTGTGGTGGGAG 1044

RESULT 11
AAT676769
ID AAT676769 standard; cDNA; 2043 BP.
AC AAT676769;
XX
DT 15-SEP-1997 (first entry)
XX
DE Human alpha 1,3/4 fucosyltransferase cDNA.
XX
KW GDP-L-fucose:beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase;
KW human; alpha 1,2 F; transgene; transgenic mouse; animal model;
KW intestinal adhesion; Helicobacter pylori infection; stomach;
KW small intestine; gut; epithelial cell; surface receptor;
KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
KW gastric adenocarcinoma; Lewis antigen; fucosylation; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 104..1201
FT /tag= a
XX
XX US5625124-A.
XX
XX 29-APR-1997.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX (UNIM ) UNIV WASHINGTON.
XX
XX Falk P, Gordon JT.
XX
XX WPI: 1997-258275/23.
```

```
DR P-PSDB; AAM23806.
XX
XX Animal model for Helicobacter pylori infection - comprising
PT transgenic mouse expressing human enzyme promoting intestinal
PT adhesion
XX
XX Example 1; Columns 23-26; 24pp; English.
XX
XX A claimed transgenic mouse expresses, in its intestinal epithelial
CC cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
CC fucosyltransferase (also called alpha 1,2 F) or human GDP-L-fucose:
CC beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
CC called alpha 1,3/4 F). The enzyme is expressed under the
CC control of a gut epithelial cell-specific promoter and Helicobacter
CC pylori adheres to the transgenic cells. The transgenic mouse and
CC intestinal epithelial cells from it are useful as models for screening
CC compounds for the ability to inhibit adhesion of H. pylori to gut
CC epithelial cells. The present sequence encodes human alpha 1,3/4 F
CC and was published by Kukowska-Latallo et al., Genes Dev., 4:
CC 1288-1303 (1990).
SQ
Sequence 2043 BP: 413 A; 647 C; 575 G; 408 T; 0 other;
Query Match 14.08; Score 128.2; DB 18; Length 2043;
Best Local Similarity 53.08; Pred. No. 1.6e-27;
Matches 374; Conservative 0; Mismatches 513; Indels 18; Gaps 4;
QY 106 TGCATCTCACAGACCGCTCATTTGACACAAATCCATGCGGTCTGATTCACCAT 165
DB 343 TGCACATCAGTCGCCGACCGCAGAGGTATACACAGCAGACCGGTATGTCACACAC 402
QY 166 AGAGACATCAGCTGAGG---TCTGACTAATCTACTCTACACAGCCGAGCCACCTTTAG 222
DB 403 TGGATATCATGTTCACACCTTAAGTACAGCGCTCCACTTCCCGAGCCGACAGGGCAG 462
QY 223 AATGATTTGGATGATTTAGTACACACCCACACACCCCAACCAAGTGGCATTTGAA 282
DB 463 CCGTGATCTGGTTCACCTGAGAGCCACCCCTAATCTGACACCTTGAAGCCCTGAGAC 522
QY 283 CACTTGTACACCTGACCTCACTTATGCGCGGATTCAGATATCCAAAGTCTTATGCG 342
DB 523 AGATACCTTCAATCTCACTGCTTCCAGCAGGAGCTCCGATCTTCAAGCCTTACGCG 582
QY 343 TTCTTGACGGTGAGCACAATCCCTTTG-----TGTTCGAAGTGCACGAAGAG 393
DB 583 TGGCTGGAGCCGTGGTCCGGCCAGCCTGCCACCACCGCTCAACCTCTGGCCAGAAC 642
QY 394 AAGTTGGTGTGCTGGTGTGTAATCTGAGTACGTAACCTGAGCATGCGAGGTTCAATATAC 453
DB 643 GAGCTGGTGGCTGGGGGGGTGTCACAACTGGAAGCGGACCTCAGCCAGGTGGCTATCTAC 702
QY 454 AACGAGCTCAGCAAGAGATTTGAATCCACACCTATGAGCCAGCTTGGGAAATACG 513
DB 703 CAGAGCTGCAGGCTCATCTCAAGTGTGAGCTGTACGAGCGCTCCACA---AACCCTG 759
QY 514 AACGATAAAATCTGATTCGCCACATCTACTTGTAAATTTTATCTTCAATTTGAAAC 573
DB 760 CCCAAGGGGACCATGATGAGAGCGTGTCCCGGTACAAAGTTTACTTGGCTTGGAGAAC 819
QY 574 TCAATTCACAAAGATTACATACAGAAAGCTCT---ACAATGCAATTTTGGCTGTTCA 630
DB 820 TCCCTGACACCCCGACTACATCACCGAAGGCTGAGAGAAAGCCCTGGAGGCTGGGCG 879
QY 631 GTACCTGTGTGCTGGTGTGCTGATCTAGGAAACCTATGAGAAATATATTCAGCTGATTC 690
DB 880 GTGCCCGTGGTGGTGGGCGCCAGCAGAAAGCACTACGAGAGGTCTCTCCACCCGAGGCC 939
QY 691 TTCATTCATGTGGAAGATTTTAATCTCTCCAGTGAATGACAAAATCTGAGCAAGT 750
DB 940 TTCACTCCAGCTGAGCAGCTTCAGAGCCCAAGAGCCTGGCCGAGTCTGAGAGAGCTG 999
QY 751 GACAAAAACAATAAGTGTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 795
```


Db 790 CAGAGCCTGCAGGCTCATCTCAAGTGTGAGCTGTACGAGCGTCCACACA---AGCCCTG 846
 QY 514 AAGGATTAATAATCTGATTTCCACCATTCTACTTGAATTTATCTTCAATTTGAAGAAC 573
 Db 847 CCCAAGGAGACCATATGATGAGCGCTGTCCGGTACAAAGTTATCTGTGGCTTCGAGAAC 906
 QY 574 TCATTTCAACAATTAATATCATCAGAAAAGCTCT---ACAATGATTTTGGCTGTGATCA 630
 Db 907 TCCCTTCACCCCGACATCACTACACCGAGAACGTGTGAGAGAACCCCTGAGAGCGCTGAGCC 966
 QY 631 GTACCTGTTGTCTGCTGCTCATCTAGGAGAAACTATGAGAAATTAATTCAGCTGATTTCA 690
 Db 967 GTCCCGCTGTGTGTGTGGCCCGACAGAGAACATAGAGAGTTCTGTCCGCCCGACGCC 1026
 QY 691 TTCAATTCATGTGAAGATTTTAACTCTCCAGTGTAGTACCAAAATATCTGAAGGAAGTT 750
 Db 1027 TTCTACACCTGTGATGACTTCCAGAGCCCGCAAGACCTGTGCGGTGACGAGACTG 1086
 QY 751 GACAAAACAAATAAGTTGTACCTTACTTACTTTACTGTGAGAAAG 795
 Db 1087 GACAAAGACACCGCCGCTACCTGAGCTACTTTCGTGGCGGAG 1131

RESULT 15

AA056905 standard; DNA: 1654 BP.

AC AA056905;
 DT 26-JUL-1994 (first entry)
 DE PCDNA1-alpha-(1-3)Fuc-TV1 genomic DNA.
 XX Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
 KM surface; oligosaccharide; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 128..1208
 FT /*tag= a
 PN MO9402616-A.
 XX 03-FEB-1994.
 PD 20-JUL-1993; 93MO-US06703.
 PF 20-JUL-1992; 92US-0914281.
 PR 20-JUL-1992; 92US-0914281.
 XX (UNMI) UNITV MICHIGAN.
 PA Lowe JB;
 PI WPI, 1994-048874/06.
 DR P-PSDB; AAR45933.
 XX DNA fragment encoding a glycosyltransferase - can be used for in
 PT vitro reactions to modify cell surface oligosaccharide(s) e.g
 PT blood gp. determinants, to protect against transplant rejection
 XX Claim 3; Fig 7; 249pp; English.
 CC The sequence is that encoding human GDP-Fuc: [beta-D-Gal (1-4)]-D-
 CC GlucNAc alpha (1-3)-fucosyltransferase. The enzyme produced by the
 CC DNA may be non glycosylated. This prevents premature loss of enzyme
 CC activity. It can also be used in in vitro reactions to modify cell
 CC surface oligosaccharide mols. e.g. blood group determinants.
 CC See also AA056906-12.
 SO Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match

13.6%; Score 124.6; DB 15; Length 1654;

Best Local Similarity 52.1%; Pred. No. 1.7e-26;
 Matches 408; Conservative 0; Mismatches 354; Indels 21; Gaps 5;

QY 31 ATTCTGTTGGGTATGCGCATTTGGGACAGACCTTTACCTTACATCTGCGCAAGCATG 90
 Db 318 ATCTGCTGTGAGAGTGTGCTTTTAAACAAACCCATACCTGTGCGCTGTGAGATG 377
 QY 91 TTCAT---ATCCAGGGTGCATCTGACACAGACCGCTCATTTGTACCAATCCCAT 147
 Db 378 GTGCTGTGACAGCGGTGATCAACATCATGCGCAGCGAAGGTGTTCACAGAGAC 437
 QY 148 GCGGTCTGATTTACCATTAAGACATACGTGGATGTGA---CTACTTACCTACAGAG 204
 Db 438 GCGGTATGCTGACACCGAGAGTGTATGTACAAACCCAGTGTCCAGCTCC 497
 QY 205 GCCAGGCGACCTTTTGAAGATGATTTGATGATTAAGTATACCCAGTACACACCC 264
 Db 498 CCGAGGCGGAGGAGGCGAGCATGTGATGTGTTGACATGAGATGCCAAGCCACTGCTGG 557
 QY 265 CAAAAGAGTGGCATTTGAACACTTGTTCACCTGACTTAACTATTCGCGTATTCAGAT 324
 Db 558 CAGCTGAAGCCCATGACGAGATATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 617
 QY 325 ATCAAGTGCCTTATGCTCTTGTACGCTGAGCAACAATCCCTTTG-----TGTTC 375
 Db 618 ATCTTCAACGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
 QY 376 GAAAGTCCCAAGCAAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
 Db 678 AACCTCTGCGCCAGAGACCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 QY 436 GCCAGGCTCAAGTATTTACAAAGAGCTGACGAGAGATTTGAAATTCACACTTATGCGCAA 495
 Db 738 GCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794
 QY 496 GCATTTGGAGATATGCTGAAAGATTAATCTGATTTCCACCATATCTTGTAAATTT 555
 Db 795 CGCTCCCAAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
 QY 556 TATCTTCAATTTGAAACCTCAATTCACAAAGATTTACATCAAGAAAGCTCTAC---AAT 612
 Db 855 TATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914
 QY 613 GCATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
 Db 915 GCCGTGAGGCGTGGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 974
 QY 673 TATATTCAGCTGATTTATTCATTCATGTGGAAGATTTTAACTCTCCAGTATGAGA 732
 Db 975 TTCTGTCACCGGACGCTTATTCACGTGAGACTTCCAGAGCCCAAGGACTGCGC 1034
 QY 733 AATATCTGAAGAGTGTGCAAAACAAATTAAGTGTACCTTGTGTTAATCTGAGAGA 792
 Db 1035 CGGTACCTGAGAGCTGTGCAAGAGACACCGCCGCTACTGAGCTACTTTCGTGGCGG 1094
 QY 793 AAG 795
 Db 1095 GAG 1097

Search completed: October 5, 2002, 23:10:04
 Job time: 121.682 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 766.105 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-3_COPY_280_1194

Perfect score: 915

Sequence: 1 acaaaactgattatttaa.....agaaatglttggaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	661.8	72.3	697	9	BB397507 BB397507
2	580.8	63.5	607	10	BI102760 BI102760
3	368	40.2	649	9	AV338075 AV338075
4	366	40.0	716	9	AV725071 AV725071
5	291.2	31.8	1101	12	CNS05957 CNS05957
6	272.2	29.7	579	9	BM021081 BM021081
7	179.6	19.6	1100	12	CNS05792 CNS05792
8	175.6	19.2	844	12	CNS01V8V CNS01V8V
9	171.4	18.7	570	10	BM082484 BM082484
10	170.8	18.7	688	9	BS95164 BS95164
11	164.8	18.0	887	12	CNS02005 CNS02005
12	156.8	17.1	1016	12	CNS0429P CNS0429P
13	152.2	16.6	199	9	BB398103 BB398103
14	150	16.4	625	9	BB649967 BB649967
15	148.4	16.2	666	10	BT050300 BT050300
16	145.8	15.9	651	10	BT036855 BT036855
17	142	15.5	640	10	BT073227 BT073227

C	18	140.4	15.3	711	12	CNS034XF	AL228012 Tetraodon
C	19	140.4	15.3	767	9	BB640911	BB640911 BB640911
C	20	133.4	14.6	501	12	BA1073	BA1073 HS-1052-B2-
C	21	118.6	13.0	674	10	BE605503	BE605503 f110h04.x
C	22	115.6	12.6	971	12	CNS047IV	AL278032 Tetraodon
C	23	110.8	12.1	1023	12	CNS02ANG	AL188773 Tetraodon
C	24	110.2	12.0	615	10	BI326811	BI326811 STR0078
C	25	106.2	11.6	487	10	BE663396	BE663396 UT-M-BH0-
C	26	106.2	11.6	720	10	BF180380	BF180380 601804564
C	27	105.8	11.6	702	10	BT043578	BT043578 BT043578
C	28	90.8	9.9	628	9	BB651409	BB651409 BB651409
C	29	90	9.8	863	12	CNS02ROF	AL210840 Tetraodon
C	30	88.4	9.7	551	10	BT035701	BT035701 BT035701
C	31	87.8	9.6	566	10	BT090565	BT090565 BT090565
C	32	85.4	9.3	601	10	BT036679	BT036679 BT036679
C	33	85.4	9.3	605	10	BT028804	BT028804 BT028804
C	34	84.8	9.3	776	12	CNS03R10	AL257289 Tetraodon
C	35	83.8	9.2	1056	12	CNS01VL7	AL169252 Tetraodon
C	36	82.6	9.0	638	10	BT035697	BT035697 BT035697
C	37	81.8	8.9	655	10	BS38740	BS38740 602566977
C	38	81.4	8.9	878	12	CNS02V1D	AL215194 Tetraodon
C	39	80.8	8.8	304	9	AU081478	AU081478 AU081478
C	40	80.8	8.8	1057	12	CNS03ADC	AL235065 Tetraodon
C	41	74.8	8.2	838	10	BS432556	BS432556 602500710
C	42	74.6	8.2	590	10	BM082755	BM082755 fu25e08.y
C	43	73.2	8.0	881	12	CNS03732	AL230807 Tetraodon
C	44	71.4	7.8	556	9	AU052985	AU052985 AU052985
C	45	71.2	7.8	505	10	BT029303	BT029303 BT029303

ALIGNMENTS

RESULT 1
LOCUS BB397507 697 bp mRNA linear EST 24-OCT-2001
DEFINITION BB397507 RIKEN full-length enriched, ES cells Mus musculus CDNA
clone C330004C19 3' similar to AB015426 Mus musculus Fut9 mRNA for
alpha1,3-fucosyltransferase IX, mRNA sequence.

ACCESSION BB397507.2 GI:16411808

VERSION BB397507.2

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 697)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M., and Hayashizaki,Y.

Muramatsu,M., and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

On Jul 15, 2000 this sequence version replaced gi:9216903.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

'S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaoka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

Location/Qualifiers
 1. 697

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_id="C330004C19"
 /clone_lib="RIKEN full-length enriched, ES cells"
 /cell_type="ES cells"
 /lab_host="ISOLR"
 /note="Site 1: XhoI. Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot-5.0 and subtraction to Rot-25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGGAGATCTCGAGTTAAATTAATTAATCCCCCCCCC 3']".

BASE COUNT

219 a 138 c 141 g 199 t

ORIGIN

Query Match 72.3%; Score 661.8; DB 9; Length 697;
 Best Local Similarity 99.6%; Pred. No. 8.1e-171;
 Matches 674; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 239 ATTGAGTCCACCCACACACCCCAAGAGTGCATGACACTGTTCACCTGA 298
 1 |
 Db 4 AATTGAGTCCACCCACACACCCCAAGAGTGCATGACACTGTTCACCTGA 62
 QY 299 CTCTACTTATCGCGGTGATTCAGATTCAGAGCTTATGCTCTTGAACGTGACA 358
 1 |
 Db 63 CTCTACTTATCGCGGTGATTCAGATTCAGAGCTTATGCTCTTGAACGTGACA 122
 QY 359 CAATCCCTTGTGTGGAAGTGCACAGAGAGAGTGTGCTGCGTGTGAGTA 418
 1 |
 Db 123 CAATCCCTTGTGTGGAAGTGCACAGAGAGAGTGTGCTGCGTGTGAGTA 182
 QY 419 ACTGAGACCTGAGCATGCGAGGTCAAGTATTCACAGAGCTCGCAAGAGTATGAAA 478
 1 |
 Db 183 ACTGAGACCTGAGCATGCGAGGTCAAGTATTCACAGAGCTCGCAAGAGTATGAAA 242
 QY 479 TCCACACCTATGCGCAAGATTCGAGATACGTAAGATTAATAATCGATCCACCA 538
 1 |
 Db 243 TCCACACCTATGCGCAAGATTCGAGATACGTAAGATTAATAATCGATCCACCA 302
 QY 539 TATCTACTTGAATTTTATCTTCAATTGAAAACCTCAATTCACAAGATTAACATCAG 598

|||||
 Db 303 TATCTACTTGAATTTTATCTTCAATTGAAAACCTCAATTCACAAGATTAACATCAG 362
 QY 599 AAAAGCTTCAGAGCATTTTGGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 658
 Db 363 AAAAGCTTCAGAGCATTTTGGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 422
 QY 659 AAAAGCTTCAGAGCATTTTGGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 718
 Db 423 AAAAGCTTCAGAGCATTTTGGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 482
 QY 719 CCAGTAGTTCAGAGCATTTTGGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 778
 Db 483 CCAGTAGTTCAGAGCATTTTGGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 542
 QY 779 ACTTACCTGAGAGAGATTTTACCTGTTGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 838
 Db 543 ACTTACCTGAGAGAGATTTTACCTGTTGCTGGTTCAGTACCTGTTGCTGGGTCACATCAGG 602
 QY 839 GCTGCGATGCGACCATGTAAAGGATCAAGATTAATGCTGTGTTAATTTAGAGA 898
 Db 603 GCTGCGATGCGACCATGTAAAGGATCAAGATTAATGCTGTGTTAATTTAGAGA 662
 QY 899 AATGTTTGGAAATTA 915
 Db 663 AATGTTTGGAAATTA 679

RESULT 2

B1102760 607 bp mRNA linear EST 26-JUN-2001
 B1102760
 LOCUS
 DEFINITION 60288338p1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5043609
 5', mRNA sequence.

ACCESSION B1102760
 VERSION B1102760.1 GI:14553653
 KEYWORDS
 SOURCE
 ORGANISM

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cga@nci.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM1119 Row: 1 column: 10
 high quality sequence stop: 606.

FEATURES

source

Location/Qualifiers
 1. 607

/organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_id="IMAGE:5043609"
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: Kidney; Vector: pCMV-Sport6; Site: 1. NCI; Site: 2. Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library. 1"
 BASE COUNT 179 a 149 c 121 g 157 t
 ORIGIN

Query Match 63.5%; Score 580.8; DB 10; Length 607;
 Best Local Similarity 99.2%; Pred. No. 1.3e-148;
 Matches 604; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Db 2 ATCAACTTGAATTTATTTATTTCTTCATTTTAAATCACTATTCCTTAAAGATTTACATCAGAG
QY 600 AAAGCTCTCAATGATTTTGGCTTCACTGTCAGTACGCTGTGTCTGGCTCCATCTAGGA
Db 62 AAAGCTCTCAATGATTTTGGCTTCACTGTCAGTACGCTGTGTCTGGCTCCATCTAGGA
QY 660 AAACATATGAAATTTATTTCTCAAGCTGATCTATTCATCTCATGTGAGAAATTTTAACTCTCC
Db 122 AAACATATGAAATTTATTTCTCAAGCTGATCTATTCATCTCATGTGAGAAATTTTAACTCTCC
QY 720 CAGTGATTTAGCAAAATTTATTCACCAAGAACTTACAAAAATCATATGAGATTTTAACTCTCC
Db 182 CAGTGATTTAGCAAAATTTATTCACCAAGAACTTACAAAAATCATATGAGATTTTAACTCTCC
QY 780 CTTTAACTTGAAGAAAGATTTTACTGTAAACCTTACCAAGCTTTTGGGAATCACTCCAG
Db 242 CTTTAACTTGAAGAAAGATTTTACTGTGTAACTTACCAAGCTTTTGGGAATCACTCCAG
QY 840 CCTGGCATGACCAACATCTAAATTTAAAGCATCAAGAAATTAAGCTCTGTGTAATTTAGAGA
Db 302 CCTGGCATGACCAACATCTAAATTTAAAGCATCAAGAAATTAAGCTCTGTGTAATTTAGAGA
QY 900 ATGGCTTTTGGAAATTTAA 915
Db 362 ATGGCTTTTGGAAATTTAA 377

REFERENCE	TITLE	JOURNAL	COMMENT
Mammalia: Eutheria; Chordata: Craniata: Vertebrata; Euteleostomi: L (basses 1 to 716)	Chen, D., Chen, Z., and Han, Z.	Unpublished (2000)	
Gu, Y., Peng, Y., Song, H., Qian, B., Liu, F., Huang, Q., Yang, Y., Gao, X., Xiao, H., Xu, X., Gu, W., Tu, Y., Jia, D., Qiu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Chen, D., Chen, Z., and Han, Z.	Unpublished (2000)		

201203 P. Housheng Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
Tel: 86-21-508091919 (ex. 45)
Fax: 86-21-508091922
Email: hanzgehc@sh.cn
This clone is for sale

	Query Match	Score	DB	Length
Best Local Similarity	40.0%	366	9	716
Matches	91.7%			
Conservative	0			
Mismatches	35			
Indels	0			
Gaps	0			

[illegible][illegible]


```

/clone="045p22"
/clone_11b="A"
/cnote="Genoscope sequence ID : COMA045DH1A1-end : 73"
BASE COUNT      287 a      249 c      269 g      278 t      18 others
ORIGIN
Query Match      31.8%; Score 291.2; DB 12; Length 1101;
Best Local Similarity 59.3%; Pred. No. 5e-69;
Matches 544; Conservative 0; Mismatches 369; Indels 5; Gaps 3;

QY      2 CAAAACCTGATTTATTTAAGAAATCTACATCTGGTGGTATGATTTGGGAGAGA 61
DB      951 CAACCAAGAGTGAATAAAATGACATTAATGCTGCTGGCTGGCCGTCGGACAAA 892
QY      62 CCTTGTGACCTTACATCTCGCAAGCAATGTTCATATCAAGGGTGCATCTCAACAG 121
DB      891 CATTGACACGTGAAGGTGTGAAGCCTTATTTGGCCTTTGAGGGGTGTATACATTAACGGCG 832
QY      122 ACCGCTCATTTGTACAAACAAATCCCATGCGGTCTGTGATTCACATAGACATGAGTGGG 181
DB      831 ACAGAAACCTTTTAAACATGCTGATGCGCTGTGCTTCCACACAGACATGCGCCGGTG 772
QY      182 ATCTGACTATCTTACCTACGACGACGACGACCCCTTTCAGAAATGATTTGGTGAAT 241
DB      771 ACCTGTCCACATGCCGACAGCTGTGCGACCTGCTCCGAAATGATCTGCTTTACA 712
QY      242 TAGAGTACCCACTACACACCCCAAAAGATGGCATTTGAACACTTGTCAACCTGACTC 301
DB      711 TGGAGTCTCCCTCAGACTCTCTAGGCACTCTGGAATGAGAGCCTGTCACTGACGCG 652
QY      302 TAACTTATCCCGTGAATTCAGATATCCAACTGCTTATGCTTCTGACGTCAGCCAA 361
DB      651 TGACATATCTGTAGATTTGACATTTCCACTGCTTATTTATCCGTTTGGCCACTG 592
QY      362 ATCCCTTTGTG---TTTGAAGTGGCAAGAGAGTGGTGGTGGTGGTGAATA 418
DB      591 ATGCTGAGGAGACTTTGTCCCACTTACGAAACAACTGAGTTTGTGATGTGTACGCA 532
QY      419 ACTGGAACCTTGAGCATGTCAGGTCATGATTTACAAAGAGCTCAGCAAGATATGAAA 478
DB      531 ACTGGAACGAGCAAGATGCAAGATTTATTAACAGATTTGTACAGATTTGTAGG 472
QY      479 TCCACACTTATGCGCAAGCATTTGGAGATGATGAAATGAAATCTGATTTCCACCA 538
DB      471 TCCACAGATTTGAGCAGCATTTGGGCAACATTTCCAGATGAAAGAACCTTACCAACA 412
QY      539 TATCTACTTAAATTTATCTTTCATTTGAAAACATCAATCAAGATTAACATCACAG 598
DB      411 TCGCCACTGCAAGTTCTACTGTCATTTGAAGATCTCATCCACAAAGACTACTTTACTG 352
QY      599 AAAAGCTCTACAAATGCAATTTTGGCTGGTTCAGTACTGTTGTCCTGGTCCATCTAGG 658
DB      351 AAAAAGCTCAACCCACTGCTGGTGGGAGACATGCGCATGCTGGGCCCACTGGGG 292
QY      659 AAAAGCTGGAATTTATTTCCAGCTGATTCATTTCAATGTCGAGAGATTTAACTGTC 718
DB      291 AGAAGCTAGGAGCAATCATCTACGAGATGCTTTATCCACGAGAGACTTGGCCAG 232
QY      719 CCAAGTGAAGCAAAATATCTGAAGAAATGACAAACAAATTAAGTTGATCTTACTGT 778
DB      231 CCAAGGAGCTTGTGATCTTACTGCTCTGCAAAAATGAGAAATTAATCTTAAAT 172
QY      779 ACTTAACTGAGAAAGATTTTACTGTAACCTACACGAGTTTGGGAATTCACATGAT 838
DB      171 ATTTTCAAGTGGGAGAGACATTTAAAGCA--GGGCAAGTGGCATTTGCCGAGGCCACT 113
QY      839 GCGTGCATGCGACCATGTAAAAAGCATCAAGATAATA--AGTCTGTGTTAATTTAGAG 897
DB      112 GTAGAGGCTTGTGATTTAGCTCAGGAGGCTNANAAATACAGAGATTCAGTCACTTANAC 53
QY      898 AAATGTTTGAATTA 915
DB      52 AAGTGTGACTGGGGTTGA 35

```

```

RESULT 6
BM021081/c
LOCUS
DEFINITION
BM021081      579 bp      mRNA      linear      EST 30-OCT-2001
1e75f03.x1 Melton Normalized Human Islet 4 M-HIS 1 Homo sapiens
cDNA 3' similar to TR:Q9Y231 Q9Y231 ALPHA-3-PUCOSYLTRANSFERASE. ;
mRNA sequence.
ACCESSION
BM021081
VERSION
BM021081.1 GI:16535437
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 579)
Melton,D., Brown,T., Kenty,G., Pernutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scaerce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bilstein,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvill,I., Williams,T.,
Jackson,Y., and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 436.
FEATURES
location/Qualifiers
1..579
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="Melton Normalized Human Islet 4 M-HIS 1"
/sex="Both"
/tissue.type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site:1: Not 1;
Site:2: Sal 1; Starting library constructed using
Superscript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
applied once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
BASE COUNT      203 a      103 c      93 g      180 t
ORIGIN
Query Match      29.7%; Score 272.2; DB 10; Length 579;
Best Local Similarity 92.6%; Pred. No. 6.8e-64;
Matches 286; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY      607 TACAATGATTTTGGCTGCTGAGTACCTGTTGCTCGGCTCAGCTAGAGAAACTAT 666
DB      579 TACAATGCTTTTCGGCTGCTGATACCTGTTGCTGGACACATGTAGGAAACTAT 520
QY      667 GAGATATATTTCCAGCTGATTCATTCATTCATGTCGAGACATTTTAACCTCCAGTGAG 726

```

Db 519 GAGATTATATTCACAGATTTCATTCATGTGAGATATTAACCTCCAGTAG 460
Oy 727 TTACCAAAATATCTGAGAGAGTTGACAAAACATTAAGTTGACCTTACCTTAAC 786
Db 459 CTACCAAGATATCTGAGAGAGTTGACAAAACATTAAGTTGACCTTACCTTAAC 400
Oy 787 TGAGAAAGAGATTTTACCTTAACCTACAGGTTTGGAGATCAGATGCTGCA 846
Db 399 TGAGAGAGAGATTTTACCTTAACCTACAGGTTTGGAGATCAGATGCTGCT 340
Oy 847 TGACACCATTAATAAGAGATCAAGATATTAAGCTGTGTGATTAATAGAAAGGTTT 906
Db 339 TGAGATCATGTGAAAGAGATCAAGATATTAAGCTGTGTGATTAATAGAAAGGTTT 280
Oy 907 TGGAATTA 915
Db 279 TGGAATTA 271

RESULT 7
CNS05792/c 1100 bp DNA linear GSS 26-JUL-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 007H05 of library A from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL324368
VERSION 1
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 1100)
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL, MEDLINE
20296633
2 (bases 1 to 1100)
Crolius,H.R., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Bouneau,L., Billault,A., Quetier,F., Saurin,W.,
Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
3 (bases 1 to 1100)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1100
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="007H05"
/clone_11b="A"
/note="Genoscope sequence ID : COAA007CD03C1-end : T7"
BASE COUNT 314 a 228 c 252 g 304 t 2 others
ORIGIN
Query Match 19.6%; Score 179.6; DB 12; Length 1100;
Best Local Similarity 55.6%; Pred. No. 2,1e-38;
Matches 430; Conservative 0; Mismatches 331; Indels 13; Gaps 4;
Oy 30 CATCTGTTGGGTATGGCCATTGGGACAGACCTTGACCTTACCTCCTCCAAAGCAAT 89

Db 774 CATCTGTTATGGTTCTGGCCGTATGACAGAGAGTTGATACGACTCTTGAAGTTTA 715
Oy 90 GTTCAATATCCAGAGGTGGCCATCTACACAGACCCGCTGATGTACCAATCCATGC 149
Db 714 CTACAAATATGACGGCTGAGACTACAGATGATGATCCCTTACGACAAAGGCGAGCG 655
Oy 150 GGTCCTATTCACATAGAGATCAGTGGAGATCTGAGTACTTACCTTACCTGACAGGCGAG 209
Db 654 TGTTCTGTCTTCTCCACAGAGATACAGTGGAAATCTGGAAACTTCCAGTGGAGCGCG 595
Oy 210 GGCACCCCTTCAGAAATGATGATGATTAAGTACCTACCTACCTACCCACCCCAAA 269
Db 594 GCGATATTTTCAGAGGTGATTTGGTTCTTACCTGGAATCTCCAGAGACAGATCAGAT 535
Oy 270 GAGTGCATGAAACACTGTGTAACCTGACTCTTAACCTTATGCGCGTATGATATCA 329
Db 534 ACCAGGCTGGAAACAGTTTTCACATGACTGTAACCTGAAAGAAATTCGACATTTGT 475
Oy 330 AGTCCCTTA---TGCTCTTGTAGAGGTGAGACAAATCCCTTGTGTAAGTCCAG 386
Db 474 AGCCAGGTACCCCTTACCATCAGAGAGAAATCTGACGAGAAATAGTTCTGCCGA 415
Oy 387 CAGAGAGAGTTGTTGCTGGTTGTGAGTACTGGAACCTGAG-----CATGCCAG 440
Db 414 GAAAAACAACTTGCCTGTGTGATGCTTACGATGCTACATCCACTGGGACAGTACAG 355
Oy 441 GGTCAAGTATTAACAGAGCTCAGCAGAGATTTGAATTCACACCTATGGCAAGCAT 500
Db 354 AGGCAATTTTATTAATGAACTTCCAAACATTAATCAATTAAGTTTGGTGGTGA 295
Oy 501 CGGAGATATGATGAAATGATTAATTCGAT---CCACCATATCTACTTGTAAATTTTA 557
Db 294 CACAGGGGTGAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 235
Oy 558 TCTTCTTTGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 617
Db 234 CTATCATTTGAGAACTGATTTACCAAAAGATATGATATGATGATGATGATGATGAT 175
Oy 618 TTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677
Db 174 TGAG 115
Oy 678 TCACAGCTGATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 737
Db 114 CCTTCAGACTCCTTATTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 55
Oy 738 TCTGAGAGAGATGACAA-AAACATTAAGTTGATGATGATGATGATGATGATGATGAT 790
Db 54 TCTGCTCTCTCTTGGACAGAGNTGAGGAGCATATGCGCTACTTGTGAGGA 1

RESULT 8
CNS01V8V 844 bp DNA linear GSS 12-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-01 end of clone
DEFINITION 198K09 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL168808
VERSION 1
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 844)
Roest Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis

us-09-744-748-3_copy-280_1194.rst

[illegible]

REFERENCE AUTHORS	TITLE	JOURNAL	COMMENT
Udammaia, Eutheria, Ara (bases 1 to 648)	Chocodara; Cernatara; Certebrata; Eurelooscomi; Kodentia; Scilographachi; Muridae; Muridae; Mus		
Hiata, T., Carniciu, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, P., Kozaki, Y., Matsuyama, T., Ito, M., Kawai, P., Konno, H., Kouta, D., Shibata, S., Shihagawa, A., Sakai, C., Sakai, K., Oino, M., Tagami, M., Tagawa, T., Takahashi, F., Takeda, Y., Tanaka, I., Toyu, T., Kurumatsu, M. and Pikien Mouse Eggs			
Unpublished (2001)			

Contract Yoshida. This sequence version replaced g1.11491766.
Laboratory for Genome Hayashizaki
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama
Fax: 81-45-503-9222
E-mail: genome-res@sc.riken.go.jp,
Carninci.P., genome.gsc.riken.go.jp/
M., Konno, H., Okazaki, Y., Muramatsu, N., Sugahara, Y., Shibata, K.,
No, Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1611-1630 (2000)
Kawahiki, M., Yoneeda, Y., Inoue, K., Togawa, I., Iwata, Y.,
S., Kawai, J., Okazaki, K., Muramatsu, M., Inoue, Y., Kita, A. and
Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a

```

/strain="C57BL/6J"
/db_xref="taxdb/6J"
/clone="C03000140090"
/clone_idb="Riken"
/stratum="Riken full-length enriched, adult male corpus
/sex="male"
/tissue_type="corpus striatum"
/dev_stage="adult"
/lab_host="PHIOB"
/notes="Site.1: Salt; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science in Riken
RIKEN, Division of Experimental Animal Research in
primed to prepare mouse tissues. 1st strand cDNA was
GAGGAGAGAGAGATCCAGACGACGCTCTTTTTTTTTTTTTTNN 3'/, cDNA was
prepared by using trehalose thermo-activated reverse
capitriptase. cDNA went through one round of normalization
to Rot 10.0 and subtraction to Rot - 195.0. Second
strand cDNA was prepared with to Rot - 195.0. Second
sequence 15' GAGGAGAGATTCCTC
3'/, cDNA was cloned into the primer adaptor
vector; a modified pluscripte KS(+) after blunt
from lambda FLIC I. Cloning sites, 5' end: Salt; 3' end:
BamHI"

```

RESULT 11	LOCUS	DEFINITION	887 bp	DNA	linear	GSS 14-MAY-2000
CNS02005/c		CNS02005				
		Tetradodon nigroviridis genome survey sequence. 1561,20 of library G from Tetradodon nigroviridis, genomic survey				
ACCESSION		AL208693				
VERSION		AL208693.1				
KEYWORDS		GSS; genome survey sequence.				
SOURCE		Tetradodon nigroviridis				

ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 887)
Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 887)
Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.

AUTHORS Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

TITLE Unpublished

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 887)
Genoscope.

AUTHORS Direct Submission

JOURNAL Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source location/Qualifiers
1..887
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_11b="G"
/note="Genoscope sequence ID : COAG156DF10LP1-end : 77"

BASE COUNT 232 a 188 c 206 g 253 t 8 others

ORIGIN

Query Match 18.0%; Score 164.8; DB 12; Length 887;
Best Local Similarity 54.8%; Pred. No. 2.3e-34;
Matches 385; Conservative 5; Mismatches 297; Indels 15; Gaps 3;

OY 30 CATTTGGTTGGGTAATGGCCATTGGGCAGACCTTTGACTTACATCTGCCAAGCAAT 89
DB 704 CATCTGTATGTTCTGTGCGGTATGACAGAGTTGGATCGATCGATCTTTGAAGATTGA 645

OY 90 GTTCATATTCACAGGGTGCATCTCAACAAGACCGCTCATTTGACAAATFCCATGC 149
DB 644 CKAATATTTGACAGCTGTGAGCTSACAGATGATGATCCTTGACGACAAAGCGCAGGC 585

OY 150 GGTCTGATTCACATAGACATCAGCTGGGATCTGACTTAATCTTACCTCAGACGGCCAG 209
DB 584 TGTCTGTTCTCCACAAAGACATACAGTGGAACTGGCAAACTGCCAGGGAGACCCGG 525

OY 210 GCCACCTTTCAGAATGATTTGATTAATAGATCACCCTACACCCCCCAAAA 269
DB 534 GCCAATATTTACAGATGATTTGTTCTTACCTGGAATCTCCAGAACACATCAGAT 465

OY 270 GAGTGCATTTGAACACTTGTCAACCTGACTTACTTATGCGCGTGAATTCAGATATCCA 329
DB 464 ACCAGCGCTGGAAACAGTTTTCACATATCTGTGACTACAGAAAGATTCGACATTTGT 405

OY 330 AGTGCCTTA---TGGCTTCTTGACGCTGAGACAAATCCCTTTGTGTTGAATGCGCAAG 386
DB 404 GGCCAGGTACCCCTTTACCATCAGAGAGAAAGACTGACGGAAATAGKTCCTCCGGA 345

OY 387 CAAAGAGAAAGTTGGTGTCTGGTGTGATGATGAAACCTTGACATGCCAGGCTCA- 445
DB 344 GAAAAACAAATTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 285

OY 446 -----AGTATTCACAGAGCTCAGCAAGAGTATTTGAATCCACACTATGCGCAAGC 497
DB 284 AAGGAAACATTTCTTTGGTGAATTCACAGCATATTCATGTTGATGTTTGGTGGTGC 225

OY 498 ATTCGAGAAATACGTGACGATAAAAATCTGATT---CCACCATATCTACTTGTAAAT 554
DB 224 GGACACAGAGAGTAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 165

OY 555 TTATCTTTTATTTGAAACTCAATTCACAAAGATTTACATACAGAAAGCTTACAAATGC 614
DB 164 TTACTGTCTCTTGGAACTCAATACACAAAGATTCATACAGAGAGAGTCAAGGAGCC 105

OY 615 ATTTTGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
DB 104 CTTTCAGATGAGACCGCTTCTGTTGTTGGACCCCAAGACCAAGCAATGAAATGTT 45

OY 675 TATTCAGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 716
DB 44 CTTCCTTCAGACTCCTTAATTCAGATGATGATTTTCTTTC 3

RESULT 12
CNS0429P/C 1016 bp DNA linear GSS 18-MAY-2000
LOCUS
DEFINITION Tetraodon nigroviridis genome survey sequence. PUC-Orl end of clone 075P02 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL271222
VERSION AL271222.1 GI:7993186
KEYWORDS GSS; genome survey sequence.

SOURCE
ORGANISM Tetraodon nigroviridis.
Tetraodon nigroviridis.
Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 1016)
Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.

AUTHORS Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

TITLE Unpublished

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1016)
Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.

AUTHORS Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

TITLE Unpublished

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1016)
Genoscope.

AUTHORS Direct Submission

JOURNAL Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source location/Qualifiers
1..1016
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_075P02
/clone_11b="G"
/note="Genoscope sequence ID : COBG075DH01SP1-end : PUC-Orl"

BASE COUNT 246 a 216 c 253 g 298 t 3 others

ORIGIN

Query Match 17.1%; Score 156.8; DB 12; Length 1016;
Best Local Similarity 55.8%; Pred. No. 3.8e-32;
Matches 386; Conservative 0; Mismatches 292; Indels 14; Gaps 4;

OY 32 TTTCGCTTTGGTATGCGCATTTTGGGACGACCTTTGACCTTACATCTGCCAAGCAATGT 91
DB 715 TTTCGCTGTGTTGTTGGCCATTCGCGACGGTGTGATTCAGAGATGCTCCACCCACT 656

VERSION	BB649967.1	GI:16484222
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 625)	
	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,U., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,J., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,K., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.	
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, url:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujitake,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Watanuki,M., Yoneda,Y., Ishikawa,T., Okawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,K., Yamada,K. and Hayashizaki,Y. K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues. Location/Qualifiers 1. 625 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="Cl30099N17" /clone_1bp="RIKEN full-length enriched, 16 days embryo head" /sex="mixed" /tissue-type="head" /dev_stage="16 days embryo" /lab_host="DH10B" /note="Site.1: SalI; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15', GAGGACAGAGAGATCCACAGAGCGCTTTTGTGTGTGTGTGTGTVN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the	

BASE COUNT	136 a	198 c	135 g	156 t
ORIGIN				
Query Match	16.4%;	Score 150;	DB 9;	Length 625;
Best Local Similarity	100.0%;	Pred. No. 2.4e-30;		
Matches 150;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACAAAAGTGTATTTTACGAAACTACCATCTGGTTGGGTATGGCCATTGGGCGAG	60	
Db	476	ACAAAAGTGTATTTTACGAAACTACCATCTGGTTGGGTATGGCCATTGGGCGAG	535	
QY	61	ACCTTGACCTTACATCTGCGCAACGATGTCATATTCACAGGGTGGCATCTCACACA	120	
Db	536	ACCTTGACCTTACATCTGCGCAACGATGTCATATTCACAGGGTGGCATCTCACACA	595	
QY	121	GACCGCTCATGTTGTACACAAATCCCATCG	150	
Db	596	GACCGCTCATGTTGTACACAAATCCCATCG	625	
RESULT 15				
Bj050300/c	Bj050300	666 bp	mRNA	linear
LOCUS	Bj050300	NIIB Mochii	normalized Xenopus	neural library Xenopus
DEFINITION	laevis cDNA clone Xl023015 3', mRNA sequence.			
ACCESSION	Bj050300			
VERSION	Bj050300			
KEYWORDS	Bj050300.1 GI:17392988			
SOURCE	EST.			
ORGANISM	African clawed frog.			
	Xenopus laevis			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Amphibia; Batrachia; Anura; Mesobatrachia; Plipodea; Plipidae;			
	Xenopodinae; Xenopus.			
	1 (bases 1 to 666)			
	Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shih-I,T. and Kohara			
	Y.			
TITLE	Expressed genes in X. laevis embryo			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Tadasu Shin-i			
	Center For Genetic Resource Information			
	National Institute of Genetics			
	1111 Yata, Mishima, Shizuoka 411-8540, Japan			
	Tel: 81-559-81-6856			
	Fax: 81-559-81-6855			
	Email: tshini@genes.nig.ac.jp.			
FEATURES	location/genes.nig.ac.jp.			
SOURCE	1..666			
	/organism="Xenopus laevis"			
	/db_xref="taxon:8355"			
	/clone="XL023015"			
	/clone_11p="NIIB Mochii normalized Xenopus neural			
	library"			
	/tissue_type="whole embryo"			
	/dev_stage="stage 15"			
BASE COUNT	178 a	116 c	157 g	215 t
ORIGIN				
Query Match	16.2%;	Score 148.4;	DB 10;	Length 666;
Best Local Similarity	55.3%;	Pred. No. 6.8e-30;		
Matches 354;	Conservative 0;	Mismatches 276;	Indels 10;	Gaps 3;
QY	251	CAACTCACACCCCCCAAAAGAGTGCGATGGAACCTGTCAACCTGACTTAATCTATC	310	
Db	666	CAAGCAACAATCCCAACAGAGGTTATAGGAAGAACTCATCATCTGACATGTCTTACC	607	
QY	311	GCCGTGATTCAGATATCCAGTGCCTTATGCGCTTCTTACGCGTGACACAAATCCCTTTC	370	
Db	606	GTCGTGACTCGATATTTTCTACTCTCTTATGCGTGTGCTGAAAGGATATGGAAGAAA	547	

Tue Oct 8 10:18:23 2002

us-09-744-748-3_copy_280_1194.rst

Page 12

OY	371	TTTTTGAAGTGGCCAGCAAGAGAAATTTGGTGTCTGGGTGTTTGAGTAAGTGCACCGCTG	430
Db	546	ACTTCATGATGCGGCCCAAGACAAACAACTAGTAGGCTGGTAGTGACGTAACTGAAACCCAA	487
OY	431	ACCATGCAGGSGTGAAGTATTTACAAACAGCTCAGCAAGAGTATTTGAAATCCACCGTAGG	490
Db	486	ACTCCAGAAAGATCAAAATTTATGAGAGTTTAAAGCCACACCTTGCTGTGTAATATTAAG	427
OY	491	GGC---AAGCATTCGGAGAAATACGTGAACGATAAATATCTATTCCACCATATCTACTTG	548
Db	426	GCGTACACCATTTTACCGCTTCCGTGACC-----AAAACACATTGAAATCTTTTCCAGTA	372
OY	549	TAAATTTATCTTTTCATTTCATTTGAAAACTCAATTCACAAAGATTATCATACAGAAAGCTCT-	607
Db	371	CAAAATTTTATTTAGTTTGTGAAAATTCAAATTCACGAAGACATCATACAGTAAGAACTCTG	312
OY	608	--ACAATGACATTTTTCGGTGGTTCAGTACCTGTTGTCTCCGGATCCATCTAGGGAAGAACTA	665
Db	311	GCATATATGCGCTCTCAATCTGTGGGTCAACCGTTGTTTTGGGGCCCTCCCGTTCTCACTA	252
OY	666	TGAGATATTAATTCACCGCTGATTCATTCAATCTAGTGAAGATTTCMACTCTCCCGTGA	725
Db	251	TGAGCGTTTTCACCCCGCAGTCATTCATCCAGTAGATGACTTTCCAACGTGTAAAGG	192
OY	726	GTTTGCAGAAATATCTGAGGAAGTTGACAAAAACAATTAAGTTGTACCTTAGTTACTTTAA	785
Db	191	ACTTGCTGATATTCCTTAACAGACGTTGATTAAGTATGATTAATTAATCATCATCAATTTCAA	132
OY	786	CTGAGAGAAAGATTTTACTGTAAACCTTACCACCGTTTTGGAGATCAATCATGCTGCCTGGC	845
Db	131	CTGGAGAGAAACATCTAAACACCAATGACCTCGAGATTTTGGAAACCCCATTAATCTGCAAGT	72
OY	846	ATGCGACCATATGTAAAGAGCATCAAGATATATAGTCTGT	885
Db	71	ATGCAAAAGCTATAAAGACGACACCGTCTATAGGACATTTT	32

Search completed: October 6, 2002, 07:15:23
Job time : 774.105 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 18.6694 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-3_COPY_280_1194

Perfect score: 915
Sequence: 1 acaaaactgattatttaa.....agaatggtttgaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
4: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCRTUS_COMB.seq:*
6: /cgn2_6/prodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128.6	14.1	1814	2	US-08-483-151-1
2	128.6	14.1	1814	5	PCT-US96-06427-1
3	128.2	14.0	1086	1	US-07-914-281-12
4	128.2	14.0	1086	1	US-08-393-246-12
5	128.2	14.0	1086	1	US-08-525-058A-12
6	128.2	14.0	1086	2	US-08-696-731-12
7	128.2	14.0	1086	4	US-09-042-531-12
8	128.2	14.0	2043	3	US-09-063-237-2
9	128.2	14.0	2043	1	US-07-914-281-1
10	128.2	14.0	2043	1	US-08-393-246-1
11	128.2	14.0	2043	1	US-08-273-411-4
12	128.2	14.0	2043	1	US-08-525-058A-1
13	128.2	14.0	2043	2	US-08-696-731-1
14	128.2	14.0	2043	5	PCT-US91-00899-1
15	128.2	14.0	2043	5	PCT-US91-00899-1
16	126.6	13.8	1316	1	US-07-914-281-10
17	126.6	13.8	1316	1	US-08-393-246-10
18	126.6	13.8	1316	1	US-08-525-058A-10
19	126.6	13.8	1316	2	US-08-696-731-10
20	126.6	13.8	1316	4	US-09-042-531-10
21	124.6	13.6	1654	1	US-07-914-281-13
22	124.6	13.6	1654	1	US-08-393-246-13
23	124.6	13.6	1654	1	US-08-525-058A-13
24	124.6	13.6	1654	2	US-08-696-731-13
25	124.6	13.6	1654	4	US-09-042-531-13
26	120	13.1	1068	5	PCT-US91-00899-13
27	120	13.1	1215	5	PCT-US91-00899-5

28	120	13.1	1488	1	US-07-914-281-9	Sequence 9, App11
29	120	13.1	1488	1	US-08-393-246-9	Sequence 9, App11
30	120	13.1	1488	1	US-08-525-058A-9	Sequence 9, App11
31	120	13.1	1488	2	US-08-696-731-9	Sequence 9, App11
32	120	13.1	1488	4	US-09-042-531-9	Sequence 9, App11
33	120	13.1	2175	4	US-08-482-073-9	Sequence 9, App11
34	120	13.1	2861	4	US-08-482-073-10	Sequence 10, App1
35	120	13.1	3647	1	US-07-914-281-7	Sequence 7, App11
36	120	13.1	3647	1	US-08-393-246-7	Sequence 7, App11
37	120	13.1	3647	1	US-08-525-058A-7	Sequence 7, App11
38	120	13.1	3647	2	US-08-696-731-7	Sequence 7, App11
39	120	13.1	3647	4	US-09-042-531-7	Sequence 7, App11
40	120	13.1	3647	5	PCT-US91-00899-4	Sequence 4, App11
41	116.8	12.8	2134	5	US-08-483-151-3	Sequence 3, App11
42	116.8	12.8	2134	5	PCT-US96-06427-3	Sequence 3, App11
43	33.8	3.7	1730	1	US-07-817-920-1	Sequence 1, App11
44	33.8	3.7	1730	1	US-08-117-006-1	Sequence 1, App11
45	33.8	3.7	1730	1	US-08-216-594-1	Sequence 1, App11

ALIGNMENTS

RESULT 1
US-08-483-151-1
Sequence 1, Application US/08483151
Patent No. 5858752
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Holgersson, Jan
TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,151
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Leach, Karen F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/278001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1814 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-483-151-1

Query Match 14.1%; Score 128.6; DB 2; Length 1814;
Best Local Similarity 52.5%; Pred. No. 5.9e-30;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

OY 104 GGTGACATCTGACACAGACCGCTATTCACCAATGCCATGGGCTGATTCACG 163
DB 548 GCTGCGCTGTGATGTCTTACCGAGCGCTGCTACCACTGATGCTGTGCTTCACG 607

Tue Oct 8 10:18:22 2002

us-09-744-748-3_copy_280_1194.rml

Page 2

```
OY 164 ATAGAGCATCAGCTGGAGTCTGACTAACTACCTCAGAGGCGGACCCCTTTCAGA 223
DB 608 ACCGTGAGCTGCAAAACCGGCAATCTCTCCACCCCTGGACAGAGGCGGACGAGCAGC 667
OY 224 AATGATTTGGATGAATTTAGAGTACCCACTCAGACCCCGGAAAGAGTGCATTAAC 283
DB 668 CTGGGCTGAGGCTCCCAATGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 727
OY 284 ACTGTTCAACCTGACTCACTAATTCGCGCTGATTCAGATTCAGATTCAGATTCAGAT 343
DB 728 GCATCTTCAACTGGGTGCTGAGCTATCGCGCTGATTCAGATTCAGATTCAGATTCAG 787
OY 344 TCTGACGAGTGCAGCAAAATCCCTTTGTTGAAGTGCAGCAAGAGAGAGTGTGT 403
DB 788 GCTTGAGAGCTCTCTCTGGGCCC--ACATCCGCACTACGGGCGGCAAAAGAGATGCTG 844
OY 404 GCTGGGTTGGATTAATGGAACCTGAGCATGCGAGGTCAGATTCAGATTCAGATTCAG 463
DB 845 CTGGGATGATCAGCAATTTCCAGAGCGGCGAGCGTGCAGAAAGTGTACCGGCGAGCTGG 904
OY 464 GCAAGAGTATTTGAATTCACACCTATGCGCCAGATTCGAGAGATTCAGATTCAGAT 523
DB 905 CCGCTCATCTGACAGGAGTGTGTGCTGCGGCGGAGGAGCGGCGGCTATCGCTAAT 964
OY 524 ATCTGATTCACCATATCTACTTAAATTTTATCTTATTTGAAATTCAGATTCAGAT 583
DB 965 GTCTGCTGCGGCGGAGTGTGTGCTGCGGCGGAGGAGCGGCGGCTATCGCTAAT 1024
OY 584 AAGATTTACATCAGAGAAAGCTCTA--CAATGCAATTTTGGGCTGGTGTACAGCTG 640
DB 1025 GGGACTCATCTAGAGAGTGTGTGCGGCGGAGGCGGCGGCTGTGCTGAGACCGCTGG 1084
OY 641 TCCGAGTCTAGAGAAATCTGAGATTTATTCAGATTCAGATTCAGATTCAGATTCAG 700
DB 1085 CGCTGGAGCTCTCTGGGCGGAGCTGTGCTGCGGCGGAGGCGGCTGTGCTGAGACG 1144
OY 701 TGGAGATTTTAACTCTCCAGTGAAGTACG 731
DB 1145 TGGAGACTTCACTGCTGCGGCTGAACTGCG 1175

RESULT 2
PCT-US96-06427-1
: Sequence 1, Application PC/TUS9606427
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation
: TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/06427
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/483,151
: FILING DATE: 07-JUN-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Lech, Karen F.
: REGISTRATION NUMBER: 35,238
: REFERENCE/DOCKET NUMBER: 00786/278W01
: TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1814 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
PCT-US96-06427-1

Query Match 14.1%; Score 128.6; DB 5; Length 1814;
Best Local Similarity 52.5%; Prod No.5.9e-30;
Matches 331; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

OY 104 GGTGCTCATCAGACAGCGCTCATCTATTCAGCAAAATCCATGCGGTCTGATTCAC 163
DB 548 GCTGCGCTGAGTGTCTAAGCAGAGCTGCTAGCGAGCTGTGCTGTGCTGTGCTGCTG 607
OY 164 ATAGAGCATCAGCTGGAGTCTGACTAATTCAGTACAGAGCGGAGGCGGAGCTTTCAGA 223
DB 608 ACCGTGAGCTGCAAAACCGGCAATCTCTCCACCCCTGGACAGAGGCGGACGAGCAGC 667
OY 224 AATGATTTGGATGAATTTAGAGTACCCACTCAGACCCCGGAAAGAGTGCATTAAC 283
DB 668 CTGGGCTGAGGCTCCCAATGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 727
OY 284 ACTGTTCAACCTGACTCACTAATTCGCGCTGATTCAGATTCAGATTCAGATTCAGAT 343
DB 728 GCATCTTCAACTGGGTGCTGAGCTATCGCGCTGATTCAGATTCAGATTCAGATTCAG 787
OY 344 TCTGACGAGTGCAGCAAAATCCCTTTGTTGAAGTGCAGCAAGAGAGTGTGT 403
DB 788 GCTTGAGAGCTCTCTCTGGGCCC--ACATCCGCACTACGGGCGGCAAAAGAGATGCTG 844
OY 404 GCTGGGTTGGATTAATGGAACCTGAGCATGCGAGGTCAGATTCAGATTCAGATTCAG 463
DB 845 CTGGGATGATCAGCAATTTCCAGAGCGGCGAGCGTGCAGAAAGTGTACCGGCGAGCTGG 904
OY 464 GCAAGAGTATTTGAATTCACACCTATGCGCCAGATTCGAGAGATTCAGATTCAGAT 523
DB 905 CCGCTCATCTGACAGGAGTGTGTGCTGCGGCGGAGGAGCGGCGGCTATCGCTAAT 964
OY 524 ATCTGATTCACCATATCTACTTAAATTTTATCTTATTTGAAATTCAGATTCAGAT 583
DB 965 GTCTGCTGCGGCGGAGTGTGTGCTGCGGCGGAGGAGCGGCGGCTGTGCTGAGACCG 1024
OY 584 AAGATTTACATCAGAGAAAGCTCTA--CAATGCAATTTTGGGCTGGTGTACAGCTG 640
DB 1025 GGGACTCATCTAGAGAGTGTGTGCGGCGGAGGCGGCGGCTGTGCTGAGACCGCTGG 1084
OY 641 TCCGAGTCTAGAGAAATCTGAGATTTATTCAGATTCAGATTCAGATTCAGATTCAG 700
DB 1085 CGCTGGAGCTCTCTGGGCGGAGCTGTGCTGCGGCGGAGGCGGCTGTGCTGAGACG 1144
OY 701 TGGAGATTTTAACTCTCCAGTGAAGTACG 731
DB 1145 TGGAGACTTCACTGCTGCGGCTGAACTGCG 1175

RESULT 3
US-07-914-281-12
: Sequence 12, Application US/07914281
: Patent No. 5324663
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B.
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
: CORRESPONDENCE ADDRESS: 14
```


Tue Oct 8 10:18:22 2002

us-09-744-748-3_copy_280_1194.rni

Page 4

Db 331 TGGGATATCATGTCACACCCCTAAGTCACAGGCTCCACCTTCCCGGAGGCGGAGGGGCGAG 390
Qy 223 AAATGATTTGGATGATTAAGTACACCCACACACCCCAAAAGAGTGGATGAA 282
Db 391 CGGTGATCTGTTCACCTTGGAGACACCCCTTAACCTACAGAGCTGGAAGGCTGGAC 450
Qy 283 CACTGTTCACCTGACCTTAATGATGCGGATTCAGATATCAAGGCTTATGGC 342
Db 451 AGATCTTCATCTCACAGTCTCACAGGACCTCGACATCTTCACGCGCTTACGAC 510
Qy 343 TTCTGAGGTGAGCACAAATCCCTTTG-----TGTGAAGTCCCAAGCAGAGAG 393
Db 511 TGGGTGAGCGGTGTGTCCGGCAGCTCCACCCACCCCTCAACCTCTCGGCGCAAGACC 570
Qy 394 AAGTGTGCTGTGGTGTGTGAGTAACTGGAACCTGAGCATGCGAGGCTCAAGTATAC 453
Db 571 GAGCTGTGGCTGTGGCGGTGTTCACACTGGAAGCCGAGCTCAGCGAGGTGCGCTACTAC 630
Qy 454 AACGAGCTCAGCAGAGTATTTGAATCCACCTATGCGCAAGCATTTGGAAATACGTG 513
Db 631 CAGAGCTGACGCTCATCTCAAGGTGAGCTGACGAGGCTCCACACA---AGCCCTG 687
Qy 514 AACGATAAAAATCTGATTCACACATATCTACTGTAAATTTTATCTTCAATTTGAAAC 573
Db 688 CCCAAGGGGACCATGATGAGAGAGCTGCTCCGATACAGTCTACCTGCGCTTGAGAC 747
Qy 574 TCAATTCACAAAGATTACATCAGAGAAAGCTCT---ACAATGCAATTTTGGCTGTCA 630
Db 748 TCGTTGACACCCCGCATACATACAGAGAGAGCTGTGAGAGAGAGCCCTGGAGGCTGGGCG 807
Qy 631 GTACTGTGTCTGGGTCCATCTGAGAAACATGATTAATTAATTCAGCTGATTA 690
Db 808 GTGCGCTGTGTGGGCGCCAGCAGAGACAGAGGTCTCTCCACCGGAGCGC 867
Qy 691 TTCAATTCATGTGAGAAATTTTAATCTCTCCAGTGAAGTATGACAAATATCTGAGAAAT 750
Db 868 TTCAATTCAGCTGTGAGAGCTCTCCAGAGCCCTGAGGCTTACTCTACAGAGACTG 927
Qy 751 GACAAAACATTAAGTGTACCTAGTACTTAACTGAGAGAG 795
Db 928 GACAAAGACACGCGCGCTACCTGAGCTACTTTCGCTGGGAG 972

RESULT 5
; Sequence 12, Application US/08525058A
; Patent No. 5770420
; GENERAL INFORMATION:
; APPLICANT: LOME, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
; STREET: 1735 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1086 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-525-058A-12

Query Match 14.0%; Score 128.2; DB 1; Length 1086;
Best Local Similarity 53.0%; Pred. No. 66-30; Indels 18; Gaps 4;
Matches 374; Conservative 0; Mismatches 313;

Qy 106 TGGCATCTCAACAGACCGCTCATTTGTACAAAGAAATCCATGCGGTCTGATTCACCAT 165
Db 271 TGGCATATCACTCCGACCGCAGAGGTGTACCCACAGGAGACAGGCTCATGCTGCACAC 330
Qy 166 AGAGATCATGCTGGGA---TGTGACTAATCTTACCTCAGAGGCGAGCCACCTTTGAG 222
Db 331 TGGGATATCATGTCCACACCTTAAGTACAGGCTCCACCTTCCGAGGCGGAGGCGAG 390
Qy 223 AAATGATTTGGATGATTAAGTACACCCACTCAGCCCAACCCCAAAAGAGTGGATGAA 282
Db 391 CGGTGATCTGTTCACCTTGGAGACACCCCTTAACCTCCAGCCTGCAACCCCTGAGC 450
Qy 283 CACTGTTCACCTGACCTTAATGATGCGGATTCAGATATCAAGGCTTATGGC 342
Db 451 AGATCTTCATCTCACAGTCTCACAGGACCTCGACATCTTCACGCGCTTACGAGC 510
Qy 343 TTCTGAGGTGAGCACAAATCCCTTTG-----TGTGAAGTCCCAAGCAGAGAG 393
Db 511 TGGGTGAGCGGTGTGTCCGGCAGCTCCACCCACCCCTCAACCTCTCGGCGCAAGACC 570
Qy 574 TCAATTCACAAAGATTACATCAGAGAAAGCTCT---ACAATGCAATTTTGGCTGTCA 630
Db 748 TCGTTGACACCCCGCATACATACAGAGAGCTGTGAGAGAGCCCTGAGAGGCTGGGCG 807
Qy 631 GTACTGTGTCTGGGTCCATCTGAGAAACATGATTAATTAATTCAGCTGATTA 690
Db 808 GTGCGCTGTGTGGGCGCCAGCAGAGAGCTGCTCCAGAGCTTACTCTACAGAGACTG 927
Qy 691 TTCAATTCATGTGAGAAATTTTAATCTCTCCAGTGAAGTATGACAAATATCTGAGAAAT 750
Db 868 TTCAATTCAGCTGTGAGAGCTCTCCAGAGCCCTGAGGCTTACTCTACAGAGACTG 927
Qy 751 GACAAAACATTAAGTGTACCTAGTACTTAACTGAGAGAG 795
Db 928 GACAAAGACACGCGCGCTACCTGAGCTACTTTCGCTGGGAG 972

RESULT 6
; Sequence 12, Application US/08696731
; Patent No. 595347
; GENERAL INFORMATION:
; APPLICANT: LOME, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOPOLYMER, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696,731
FILING DATE: 14-AUG-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-696-731-12

Query Match 14.0%; Score 128.2; DB 2: Length 1086;
Best Local Similarity 53.0%; Pred. No. 6e-30;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

DB 106 TGGCATCTGACAGACGCGCTCATGTACACAAATCCCATGCGGCTTACACAT 165
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 271 TGGCATCTGACAGACGCGCTCATGTACACAAATCCCATGCGGCTTACACAT 330
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 166 AGAGATCATGCTGGA---TCTGACTTAACCTACCTAGAGCGGACCGACCTTTGAC 222
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 331 TGGATATCATGTCACCAACCTTAAGTACAGCGCTCCACCTTCCCGAGCGCGGAG 390
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 223 AATGATTTGATGATTTAGAGTACACGACACACACCCCAAAAGAGTGCATTGAA 282
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 391 CGCTGATCTGTTCAACTTGGAGCACCCTTAACCTGACAGACCTGGAAGCCCTGGAC 450
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 283 CACTTGTCAACCTGACTCACTTAATTCGCGGATTCAGATATCCAAAGTCCCTTGGC 342
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 451 AGATATCTCACTTCACTATGCTTACCGGAGGAGACTCCGACATCTTCACGCGCTTACG 510
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 343 TTTCTACGCTGAGCAAAATCCCTTTG-----TCTTGAAGTCCCAAGAGAG 393
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 511 TGGCTGAGAGCGCTGTCCGCGCAGCTGCCACCAACCGCTCAACCTTCGCGCAAGAGC 570
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 394 AATGATTTGATGATTTAGAGTACACGACACACACCCCAAAAGAGTGCATTGAA 453
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 571 GAGCTGTGCTGCGGCTGTCTCAACTGGAAGCCGAGTCAAGCAGGCTGCGCTACTAC 630
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

DB 454 AACGAGCTGACAGAGATATTTGAATTCACACCTATGCGCAACATTCGAGATATCTG 513
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 631 CAGAGCTGAGGCTCATCTCAAGGTGAGAGGTACGAGAGCTCCACACA---AGCCCTG 687
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 514 AACGATTAATATGATTTCCACCATATCTACTTGAATTTATCTTCATTTGAAAC 573
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 688 CCCAAGGGAGCAATGATGAGAGCTGCTCCGCTTACAGTTCTACCTGCGCTGGAAGC 747
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 574 TCATTCACAAAGATTAATCATACAGAAAGCTCT---ACAATGATTTTGGCTGTGTA 630
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 748 TCCCTTACACCCGACTACATACCGAGAAAGCTTGAGAGAGCCCTGAGAGCTTGGCC 807
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 631 GTACCTGTGCTGCTGCTCATCTAGGAAACTATGAGATTAATTCAGCTGATTTCA 690
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 808 GTCCCGTGTGCTGTGGCCCGACAGAGCAATGAGAGAGTCTTCCGACCCGAGAGC 867
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 691 TTTATTCATGTTGGAAGATTTTAACTCTCCAGTGTAGCAAAATATCTGAAGAGTT 750
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 868 TTTATTCAGTGTGAGCAGCTTCCAGAGAGCCCAAGAGCTGCGCTGAGAGAGCTG 927
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 751 GACAAAACATTAAGTTTACCTTAACTTAACTGAGAGAG 795
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 928 GACAGAGACACGCGCGCTGAGCTTACCTTCCGCTGGCGGAG 972
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

RESULT 7
US-09-042-531-12
Sequence 12, Application US/09042531
Patent No. 6268193
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOPOLYMER, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid

DB 880 GTGCCCCGTGCTGCTGGCCCCCAGACAGCACTACGAGAGTTCTTCCGCCGCCGACGCC 939
QY 691 TTCATTCATGTGAGAGATTTTAACTCTCCAGTAGTAGCAAAATATCTGAAGAATT 750
DB 940 TTCATTCAGTGGAGACTTCCAGAGCCCCAGAGACTGGCCGGTAGCTGAGAGCTG 999
QY 751 GACAAAACAAATAGTTGACTTACTTAACTGGAAGAAG 795
DB 1000 GACAAAGACCACGCCCGCTACCTGACTTCTGCTGGCGGAG 1044

RESULT 9
US-07-914-281-1
Sequence 1, Application US/07914281
Patent No. 5324663
GENERAL INFORMATION:
APPLICANT: LOME, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C.,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
ANTI-SENSE: NO
US-07-914-281-1

Query Match 14.0%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred. No. 8.3e-30;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

QY 106 TGGCATCTCACAACAGCCGCTCATTTGTACAAACAAATCCGCTGCTGATTCACAT 165
DB 343 TGGCATCTCACAACAGCCGCTCATTTGTACCAAGGAGTACCAAGGAGACAGGCTCATCGTACAC 402
QY 166 AGAGATATAGCTGGGA---TCTGACTAATCTACCTCAGACGCCAGGCCACCTTTGAG 222
DB 403 TGGGATATATATGTCACACCTTAAGTACGCTCCACCTTCCCGGCGGAGGAGGAG 462
QY 223 AATGATTTGATGATTTAGAGTACACCACTACACCCCAAAAGAGTGGCATTTGA 282
DB 463 CGGTGATCTGTTGATCACTTGGAGCCACCCCTAACTGCCAGACTGGAAGCCCTGGAG 522

QY 283 CACTTGTCACTGACTACTTAATTCGCCGTGATTCAGAAATTCAGATGCCCTTATG 342
DB 523 AGATACCTTCAATCTACCATATGCTTACCAGAGACTCCGACATCTTACGCCCTG 582
QY 343 TTCTGACGGGTGACCAAAATCCCTTG-----TGTGAACTGGCAAGCAGAG 393
DB 583 TGGCTGAGCCGTGTCGCCGCCAGCCTGCCACCCAGCCTCAACTCTGGCGCAAGACC 642
QY 394 AAGTGTGCTGCTGGGTGTGAGTAATCTGAGAACCTGAGCATCGAGGTCAAGTATAC 453
DB 643 GAGCTGTGCTGCTGGCGGGGTGTCACACTGGAACCGGACTCAGCCAGGGTGGCTACTAC 702
QY 454 AACGACTCAGCAAGATTTGAATTCACACCTATGAGCCAGAACATCGAGAAATACG 513
DB 703 CAGAGCTGAGGCTCATCTCAAGGTGAGCTGTACGAGCGTCCACACA---AGCCCTG 759
QY 514 AACGATMAAAATGATTTCCACCATATCTACTTGAATTTTATCTTATCTTATGAAAC 573
DB 760 CCCAAGGGGACCATGATGAGAGCGTGTCCGGTACAAATCTACCTGGCTTCGAGAAC 819
QY 574 TCAATTCACAAAGATTAATACATACAGAAAGCTCT---ACAATGATTTTGGCTGTTCA 630
DB 820 TCTTGACACCCGACTACATACACGAAAGCTGTGAGGAGCCCTGGAGGCTTGGCC 879
QY 631 GTACCTGTTGCTGCTGCTCATCTAGGAAACTATGAGAAATTAATTCAGCTGATTTCA 690
DB 880 GTGCCCGTGTGCTGGGCCCCAGCAGAGCACTACGAGAGGTCTGCGACCCGAGCG 939
QY 691 TTCATTCATGTGAGAGATTTTAACTCTCCAGAGATTTAGCAAAATATCTGAAGAATT 750
DB 940 TTCATTCAGTGGAGACTTTCACCTTACTTAACTGGAAGAAG 795
QY 751 GACAAAACAAATAGTTGACTTACTTAACTTAACTGGAAGAAG 795
DB 1000 GACAAAGACCACGCCCGCTACCTGACTTCTGCTGGCGGAG 1044

RESULT 10
US-08-393-246-1
Sequence 1, Application US/08393246
Patent No. 5595900
GENERAL INFORMATION:
APPLICANT: LOME, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C.,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451

Tue Oct 8 10:18:22 2002

us-09-744-748-3_copy_280_1194.rn1

Page 8

REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEFAX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ANTI-SENSE: NO
US-08-393-246-1

Query Match 14.0%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred. No. 8.3e-30;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
DB 106 TGCATCTCAACAAGACGCTCATTTAGACAATAATCCATGCGCTGATTCACAT 165
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343 TGCCACATCACTGCGCCGCAAGAGGTACCACAGGAGACAGCATGTCAGCAC 402
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 166 AGAGACATCACTGCGCA--TCTGACTAATCTACAGAGCGCGGCTGATTCACAT 222
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 403 TGGATATCATGTCCAACTTAAGTACGCTCCACCTTCCGAGGCGGAGGCGAG 462
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 223 AATGATTTGATGATTTAGATGACCCACTACACACCCCAAGAGAGTGGCATTTGA 282
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 463 CGCTGGATCTGTTCACTTGAAGGACCCCTTAACCTGACAGCTGAAAGCCCTGAGC 522
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 283 CACTTGTCAACTGACTTAACTATGCGCGTATTCAGATATTCAGATGCTTATGCG 342
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 523 AGATACCTCAATCTACATGTCCTACGCGACGACGCTCCGACATCTTCACGCGCTAC 582
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343 TTCTGACGCTGACACAAATCCCTTG-----TGTGAAAGGCCAAGCAAGAG 393
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 583 TGCGTGAAGCGGTGTGTCGCGGCGAGCGCTGCCACCGCTCAACCTCTGCGCAAGCC 642
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 394 AAGTGTGTGTGCTGGGTGTGAGTAACTGAGAACTGTGAGATGCGAGGCTCAAGTATTC 453
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 643 GAGCTGTGTGCTGGGTGTGAGTAACTGAGAACTGTGAGATGCGAGGCTCAAGTATTC 702
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 454 AAGAGCTCAGCAAGAGTATTAATCCACACTATGAGCCCAAGCATTGCGAGATACGTG 513
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 703 CAGAGCTCAGAGGCTCATCTCAAGGTGAGGTGAGGAGCGCTCCACA--AGCCCTG 759
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 514 AAGATTAATAATCTGATTCACCATATCTATTAATTTATCTTTTATTTGAAAC 573
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 760 CCCAAGGGGACATGATGAGAGCGTGTCCGGGTACAAAGTCTACGCTGCGCTGAGAAC 819
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 574 TCAATTCACAAGATTTATCAACAGAAAGCTCT--ACATGATTTTGGTGTGTTCA 630
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 820 TCTTGTGACCCGACTACATCACTGAGAGAGCTGTGAGAGAGCGCTGAGAGGCTGGGCG 879
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 631 GTACTGTGTCTGCTGCTCATCTAGGAAAATCTAGAAATTAATCTAGCTGATTTCA 690
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 880 GTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 939
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 691 TTTCAATTCATGTAAGATTTTAACTCTCCAGTGAATGATTAATTAATTTGAGAGAGTT 750
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 940 TTTCAATTCATGTAAGATTTTAACTCTCCAGTGAATGATTAATTAATTTGAGAGAGTT 750
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 751 GACAAACATAGTGTGATCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 795
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1000 GACAAACATAGTGTGATCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1044
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-08-273-411-4
Sequence 4, Application US/08273411
Patent No. 5625124
GENERAL INFORMATION:

APPLICANT: Falk, Per
APPLICANT: Gordon, Jeffrey I.
TITLE OF INVENTION: Animal Model for Gastro-Intestinal
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,411
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: WJ106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 815-6558
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..361
OTHER INFORMATION: /note= "Nucleotides 73 through 1158 encode the GDP-L-fuco
PUBLICATION INFORMATION:
AUTHORS: Kukowska-Latallo, et al.
JOURNAL: Genes & Development
VOLUME: 4
PAGES: 1288-1303
DATE: 1990
RELEVANT RESIDUES IN SEQ ID NO: 4: FROM 1 TO 2043
US-08-273-411-4
Query Match 14.0%; Score 128.2; DB 1; Length 2043;
Best Local Similarity 53.0%; Pred. No. 8.3e-30;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
DB 106 TGCATCTCAACAAGACGCTCATTTAGACAATAATCCATGCGCTGATTCACAT 165
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343 TGCCACATCACTGCGCCGCAAGAGGTACCACAGGAGACAGCATGTCAGCAC 402
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 166 AGAGACATCACTGCGCA--TCTGACTAATCTACAGAGCGCGGCTGATTCACAT 222
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 403 TGGATATCATGTCCAACTTAAGTACGCTCCACCTTCCGAGGCGGAGGCGAG 462
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 223 AATGATTTGATGATTTAGATGACCCACTACACACCCCAAGAGAGTGGCATTTGA 282
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 463 CGCTGGATCTGTTCACTTGAAGGACCCCTTAACCTGACAGCTGAAAGCCCTGAGC 522
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 283 CACTTGTCAACTGACTTAACTTATGCGCGTATTCAGATATTCAGATGCTTATGCG 342
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 523 AGATACCTCAATCTACATGTCCTACGCGACGACGCTCCGACATCTTCACGCGCTAC 582
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343 TTCTGACGCTGACACAAATCCCTTG-----TGTGAAAGGCCAAGCAAGAG 393
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 583 TGCGTGAAGCGGTGTGTCGCGGCGAGCGCTGCCACCGCTCAACCTCTGCGCAAGCC 642
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Tue Oct 8 10:18:22 2002

us-09-744-748-3_copy_280_1194.rni

Page 10

ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696,731
FILING DATE: 14-AUG-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavallie, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
STRAND: nucleic acid
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
ANTI-SENSE: NO
US-08-696-731-1

Query Match 14.0%; Score 128.2; DB 2; Length 2043;
Best Local Similarity 53.0%; Pred. No. 8.3e-30;
Matches 574; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
DB 106 TGCATCTCACACAGACCGCTCATTTGTACAAATCCATGCGTCTGATTCACCAT 165
DB 343 TGCACATCTACCGCCGAGAGTGTACCCAGAGCAGACGATCTGACACCAT 402
DB 166 AGACATCTACCGGGA---TGTACTAATCTTACCTGAGAGCCACCTTTGAC 222
DB 403 TGGATATCATCTGTCACACCTTACCTGACGCTTCCACCTTCCGAGCGGAG 462
DB 223 AATGATTTGATGATTTAGATTCACCTGACACCCGCCAAAGAGTGCATTGAA 282
DB 463 CGCTGATCTGCTTCACTTGGAGCCACCCCTTAACCTGACGACCTGGAAGCC 522
DB 283 CACTTGTTCACCTGATCTTACTTATGCGCGTATTCAGATATCCAAAGTCC 342
DB 523 AGATATCTCAATCTCACACCTTCTACCTGACGACCTTCCACATCTTACG 582
DB 343 TTTTGAAGGTGACACAAATCCCTTTG-----TGTTTGAAGTGCACAAAG 393
DB 583 TGGCTGAGCGCGTGTGCTGGGACCCGCCACCCACCTGCAACCTCTGCGCA 642
DB 394 AAGTGTGTGTGCTGGTGTGAGTAATCTGGAACCTGAGCATGCGAGGCTGAT 453
DB 643 GAGCTGTGTGCTGGGCTGCTGCTCACTGGAACCGGAGCTCACAGGCGCT 702
DB 454 AAGGAGCTCAGCAAGATATTGAATTCACCTATGAGCCCAAGGAGAGTAAG 513
DB 703 CAGAGCTGAGCGCTCATCTCAAGGTGAGGTGACGAGCGCTCCACA---AG 759
DB 514 AAGCATTTAAATCTGATTTCCACCATATCTACTTGAATTTATCTTCAATT 573
DB 760 CCAAGGAGCAATGATGAGAGCTGTCCCGGTACAGTTTACCTGCGCTTGA 819
DB 574 TCAATTCACAAAGATTCATACAGAAAGCTCT---ACAATGCAATTTTGG 630

DB 820 TCTTGCACCCGACATACATACCGAAGACTGTGGAGAACCCCTGGAGCC 879
DB 631 GTACCTGTGCTCCGATCTAGAGAAATATGAGAAATATTTGACCTATTTCA 690
DB 880 GTGCCGCTGTGCTGGGCCCCAGACAGAACTAGAGAGTCTCTGCAACCGAG 939
DB 691 TTTTATGATGAGATTTTAACTCTCCAGAGATTTGCAAAATATCTGAGAG 750
DB 940 TTTATCAGCTGAGCACTTCAGAGCCCAAGAGCTGCGGCTGAGAGAGCTG 999
DB 751 GACAAACATATAGTTTACTTACTTACTTACTTACTGAGAAAG 795
DB 1000 GACAAAGACCAAGCCGCTACCTGAGCTACTTCTGCTGGCGGAG 1044

RESULT 14
US-09-042-531-1
Sequence 1, Application US/09042531
Patent No. 6268193
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOPOLYMER OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: ORION, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavallie, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRAND: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
ANTI-SENSE: NO
US-09-042-531-1
Query Match 14.0%; Score 128.2; DB 4; Length 2043;
Best Local Similarity 53.0%; Pred. No. 8.3e-30;
Matches 574; Conservative 0; Mismatches 313; Indels 18; Gaps 4;
DB 106 TGCATCTCACACAGACCGCTCATTTGTACAAATCCATGCGTCTGATTCACCAT 165

```
Db 343 TCCGACATCAGTGGCGACGCAAGGTGATCCACAGGCGACAGGCTCATCGCACAC 402
Oy 166 AGAGACATCAGTGGGA---TCGACTAATTACCTCAGCAGGCGCCACCTTTTCAG 222
Db 403 TGGGATATATATCCACACCTTAAGTCAGGCTCCACCTTCCCGAGGCGGAGGGGCG 462
Oy 223 AATGATTTGGATGAATTTAGAGTCACCCACTCAGCAGCCCAAGAGAGTGGCATTTGA 282
Db 463 CCTGGATCTGTTCAACTTGGAGCCACCCCTTAAGTGCAGACCTGGAAGCCCTGAGC 522
Oy 283 CACTGTCTAAGCTGACTAATTATGCGCGTGAATTCAGATATCCAAATGCTTATGCG 342
Db 523 AGATACCTTAATCTCAGCATGTCTTACGCGACGACTCCGACATCTTCAAGCCCTACG 582
Oy 343 TTTCTGACGCTGAGCACAATCCCTTTG-----TGTGAAATGCGCAAGCAGAG 393
Db 583 TGGCTGAGACCGCTGCTGCGCGCCAGCCTGCCACCCCTCAACCTCTGCGCCAAAGAC 642
Oy 394 AAGTGTGTGCTGGTGTGAGTAAGTGAACCTGAGCATGCGAGGTCAGATTTAC 453
Db 643 GAGCTGTGGCTGGCGGCTGTCCAACTGGAAGCGGACTCAGCGAGGTCGCTAC 702
Oy 454 AACGAGCTGAGCAAGATTTGAATTCACACCTATGCGCAAGCATGCGGAATATCGTG 513
Db 703 CAGAGCTGACAGCTCATCTCAAGGTGAGCTGACGAGCGCTCCACA---AGCCCTG 759
Oy 514 AACGATTAATAATCTGATTCCTCCACATATCTACTGTAATTTATCTTATTTGAAAC 573
Db 760 CCCAAGGGACCATATATGAGAGCGCTGTCCGGTACAACTTCTACCTGCGCTTCGAAAC 819
Oy 574 TCAATTCACAAAGATTATCATCAGAAAAGCTCT---ACAATGCAATTTTGGCTGGTTCA 630
Db 820 TCTTGCACCCCGACTACATCATCAGGAGCTGTGGAGGAGAGCCCTGAGGCGTGGGCG 879
Oy 631 GTACCTGTGTCTGGGTGCATCTAGGGAACATATGAAATATATTTCCAGCTATTTA 690
Db 880 GTGCCGTGTGTCTGGGCGGCGGAGAGCAAGCACTACGAGAGGTCTCTGCGACCGCGC 939
Oy 691 TTTCTATCATGTGAGAAATTTAACTCTCCAGTGAAGTTAGCAAAATATCTGAAGAAAT 750
Db 940 TTATTCACAGTGTGAGACTTCCAGAGCCCAAGAGACTGCGCGGTACTGCAAGAGCTG 999
Oy 751 GACAAAACAAATAGTTGACTTACTTAACTTAACTGGAAGAAG 795
Db 1000 GACAAGAGACGAGCGCGGTACTGAGCTACTTTCGCTGCGGAG 1044
```

RESULT 15
PCT-US91-00899-1
Sequence 1, Application PC/TUS9100899

GENERAL INFORMATION:

APPLICANT: Lowe, John B.

TITLE OF INVENTION: Method and Products For the Synthesis of

TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,

TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/00899

FILING DATE: 19910214

```
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavelle Ph.D., Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-5940
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Blood
CELL LINE: A431
PCT-US91-00899-1
```

Query Match 14.0%; Score 128.2; DB 5; Length 2043;
Best Local Similarity 53.0%; Pred. No. 8,3e-30;
Matches 374; Conservative 0; Mismatches 313; Indels 18; Gaps 4;

```
Oy 106 TCCGATCTCACAACAGCCGCTATTTGTACAAACAAATCCCATGCGGTCTGATTCACAT 165
Db 343 TCCGACATCAGTGGCGACGCAAGGTGATCCACAGGCGACAGGCTCATCGCACAC 402
Oy 166 AGAGACATCAGTGGGA---TCGACTAATTACCTCAGCAGGCGCCACCTTTTCAG 222
Db 403 TGGGATATATATCCACACCTTAAGTCAGGCTCCACCTTCCCGAGGCGGAGGGGCG 462
Oy 223 AATGATTTGGATGAATTTAGAGTCACCCACTCAGCAGCCCAAGAGTGGCATTTGA 282
Db 463 CCTGGATCTGTTCAACTTGGAGCCACCCCTTAAGTGCAGACCTGGAAGCCCTGAGC 522
Oy 283 CACTGTCTAAGCTGACTAATTATGCGCGTGAATTCAGATATCCAAATGCTTATGCG 342
Db 523 AGATACCTTAATCTCAGCATGTCTTACGCGACGACTCCGACATCTTCAAGCCCTACG 582
Oy 343 TTTCTGACGCTGAGCACAATCCCTTTG-----TGTGAAATGCGCAAGCAGAG 393
Db 583 TGGCTGAGACCGCTGCTGCGCGCCAGCCTGCCACCCCTCAACCTCTGCGCCAAAGAC 642
Oy 394 AAGTGTGTGCTGGTGTGAGTAAGTGAACCTGAGCATGCGAGGTCAGATTTAC 453
Db 643 GAGCTGTGGCTGGCGGCTGTCCAACTGGAAGCGGACTCAGCGAGGTCGCTAC 702
Oy 454 AACGAGCTGAGCAAGATTTGAATTCACACCTATGCGCAAGCATGCGGAATATCGTG 513
Db 703 CAGAGCTGACAGCTCATCTCAAGGTGAGCTGACGAGCGCTCCACA---AGCCCTG 759
Oy 514 AACGATTAATAATCTGATTCCTCCACATATCTACTGTAATTTATCTTATTTGAAAC 573
Db 760 CCCAAGGGACCATATATGAGAGCGCTGTCCGGTACAACTTCTACCTGCGCTTCGAAAC 819
Oy 574 TCAATTCACAAAGATTATCATCAGAAAAGCTCT---ACAATGCAATTTTGGCTGGTTCA 630
Db 820 TCTTGCACCCCGACTACATCATCAGGAGCTGTGGAGGAGAGCCCTGAGGCGTGGGCG 879
Oy 631 GTACCTGTGTCTGGGTGCATCTAGGGAACATATGAAATATATTTCCAGCTATTTA 690
Db 880 GTGCCGTGTGTCTGGGCGGCGGAGAGCAAGCACTACGAGAGGTCTCTGCGACCGCGC 939
Oy 691 TTTCTATCATGTGAGAAATTTAACTCTCCAGTGAAGTTAGCAAAATATCTGAAGAAAT 750
Db 940 TTATTCACAGTGTGAGACTTCCAGAGCCCAAGAGACTGCGCGGTACTGCAAGAGCTG 999
Oy 751 GACAAAACAAATAGTTGACTTACTTAACTTAACTGGAAGAAG 795
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:39:45 ; Search time 3536.27 Seconds
(without alignments)
16699.717 Million cell updates/sec

Title: US-09-744-748-4

Perfect score: 2822
1 aagcttgctactcagagag.....aacatgttcgattgaattc 2822

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank1.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pac.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description

1	2814	99.7	80247	9	AL512406	AL512406 Human DNA
2	1542	54.6	3019	9	AB023021	AB023021 Homo sapi
3	1537.2	54.5	2501	9	HS238701	AU238701 Homo sapi
4	1103.2	39.1	1705	10	AF230460	AF230460 Cricetulu
5	1092.2	38.7	2139	10	AB015426	AB015426 Mus muscu
6	1085.4	38.5	2156	10	AB049819	AB049819 Rattus no
7	937.6	33.2	1108	10	AF345993	AF345993 Rattus no
8	823.2	29.2	1707	5	AB035906	AB035906 Gallus ga
9	676.8	24.0	1080	5	AB035905	AB035905 Xenopus 1
10	644	22.8	71396	2	AC100674	AC100674 Mus muscu
11	557.8	19.8	561	11	G55571	G55571 SHGC-100926
12	281	10.0	71396	2	AC100674	AC100674 Mus muscu
13	279.6	9.9	1567	5	AB023627	AB023627 Dario rer
14	234.6	8.3	1832	5	AB023628	AB023628 Dario rer
15	222.6	7.9	1092	5	AB035907	AB035907 Xenopus 1
16	220.2	7.8	1264	5	AF288369	AF288369 Gallus ga
17	176.2	6.2	2657	5	CGU73678	U73678 Gallus gall
18	145.2	5.1	1429	10	CGU78737	U78737 Cricetulu
19	145.2	5.1	1462	10	AF090449	AF090449 Cricetulu
20	143.6	5.1	1718	10	AF090449	AF090449 Cricetulu
21	137.2	4.9	1044	5	AB035908	AB035908 Gallus ga
22	132	4.7	2293	4	BTA134FTG	X87810 B. taurus al
23	132	4.7	12607	4	BTA132772	AJ132772 Bos tauru
24	127.6	4.5	1158	10	AB039126	AB039126 Mus muscu
25	127	4.5	1487	3	AF016899	AF016899 Schistos
26	127	4.5	1814	6	AR028686	AR028686 Sequence
27	127	4.5	3594	10	MMU45980	U45980 Mus muscu
28	127	4.5	117993	2	AC091284	AC091284 Mus muscu
29	125.6	4.5	1613	4	BTA132776	AJ132776 Bos tauru
30	125.6	4.5	1619	4	BTA132775	AJ132775 Bos tauru
31	125.6	4.5	1644	4	BTA132774	AJ132774 Bos tauru
32	125.6	4.5	1873	4	BTA132773	AJ132773 Bos tauru
33	124.6	4.4	1119	9	AF345881	AF345881 Macaca mu
34	122.8	4.4	1170	10	AB039127	AB039127 Mus muscu
35	122.8	4.4	1170	10	AB039131	AB039131 Mus muscu
36	122.8	4.4	1615	10	MMU34584	U34584 Mus muscu
37	122	4.3	1113	9	AF345884	AF345884 Saltili s
38	121.2	4.3	1170	10	AB039124	AB039124 Mus muscu
39	121.2	4.3	1170	10	AB039125	AB039125 Mus muscu
40	121.2	4.3	1170	10	AB039128	AB039128 Mus muscu
41	121.2	4.3	1170	10	AB039129	AB039129 Mus muscu
42	121.2	4.3	1170	10	AB039130	AB039130 Mus muscu
43	121.2	4.3	1170	10	AB039132	AB039132 Mus muscu
44	119.8	4.2	1119	9	AF345883	AF345883 Macaca mu
45	118.4	4.2	1170	10	AB039133	AB039133 Mus spicl

ALIGNMENTS

RESULT 1
LOCUS AL512406
DEFINITION Human DNA sequence from clone RPl1-5049 on chromosome 6, complete sequence.
ACCESSION AL512406
VERSION AL512406.14 GI:13897154
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 80247)
AUTHORS Bates, K.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Apr 30, 2001 this sequence version replaced gi:13396709. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, S4, SWISSPROT, Tr, TrEMBL, Wp, WORMPEP, information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> Rpl1-504J9 is from the library RPl1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/Bacpac/home.htm> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone Rpl1-504J9. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone Rpl1-381A2 is at 80148 in this sequence. The true right end of clone Rpl1-77B15 is at 100 in this sequence.

FEATURES

```
source
1..80247
/organism="Homo sapiens"
/DB_xref="taxon:9606"
/chromosome="6"
/clone="Rpl1-504J9"
/clone_lib="RPl1-11.2"
184..1422
/note="L2 repeat: matches 875..2419 of consensus"
repeat_region
1744..1834
/note="MIR repeat: matches 166..258 of consensus"
repeat_region
2950..2977
/note="14 copies 2 mer ca 100% conserved"
repeat_region
3411..3669
/note="MIR16c repeat: matches 93..364 of consensus"
4372..4704
/note="MIR1A1 repeat: matches 5..365 of consensus"
/note="match: GSS: Em:A0550748"
8126..8184
/note="MER58B repeat: matches 272..332 of consensus"
8185..8696
/note="MIR2D repeat: matches 1..513 of consensus"
8779..8817
/note="MIR2D repeat: matches 504..542 of consensus"
8819..8974
/note="MER58A repeat: matches 6..162 of consensus"
/note="match: GSS: Em:B90633"
complement(9443..9891)
/note="match: GSS: Em:A0261227"
complement(9452..9883)
/note="match: GSS: Em:A0119695"
9782..9894
/note="MIR repeat: matches 92..205 of consensus"
10089..10212
/note="MIR repeat: matches 135..262 of consensus"
complement(10093..10671)
/note="match: GSS: Em:A037061"
10223..10282
/note="MIR repeat: matches 203..262 of consensus"
complement(10315..10631)
/note="match: GSS: Em:A0200259"
10805..11186
/note="match: GSS: Em:A072584"
11657..11844
/note="L2 repeat: matches 2546..2750 of consensus"
complement(11659..112084)
misc_feature
complement(11674..11987)
/note="match: GSS: Em:A094678"
12083..12588
/note="match: GSS: Em:A0691235"
12596..12810
/note="MIR repeat: matches 23..248 of consensus"
12920..12986
/note="LIM5 repeat: matches 6228..6295 of consensus"
12987..13169
/note="LIM5 repeat: matches 5852..6043 of consensus"
13168..13280
/note="LIM5 repeat: matches 5710..5832 of consensus"
13851..17218
/note="LIM8 repeat: matches 2787..6155 of consensus"
18903..18953
/note="MIR2CB repeat: matches 439..501 of consensus"
18984..19433
/note="MIR2CB repeat: matches 1..460 of consensus"
complement(19108..19690)
/note="match: GSS: Em:A0563935"
20683..20801
/note="AluJ6/FRAM repeat: matches 169..288 of consensus"
20802..20862
/note="AluJ6/FRAM repeat: matches 3..66 of consensus"
20912..21033
/note="MIR repeat: matches 114..237 of consensus"
21047..21128
/note="41 copies 2 mer ta 63% conserved"
21676..21859
/note="L2 repeat: matches 2130..2308 of consensus"
21890..23612
/note="LIP repeat: matches 841..2570 of consensus"
23601..26609
/note="LIP repeat: matches 3107..6145 of consensus"
26620..26789
/note="LIP3 repeat: matches 1107..6145 of consensus"
26821..28627
/note="MER5A repeat: matches 1..174 of consensus"
28656..28837
/note="MER39 repeat: matches 5..111 of consensus"
28838..29142
/note="MER39 repeat: matches 54..235 of consensus"
29143..29494
/note="AluSC repeat: matches 4..308 of consensus"
30529..31344
/note="MER39B repeat: matches 235..549 of consensus"
32234..32340
/note="LIP3 repeat: matches 5331..6146 of consensus"
32399..32578
/note="L2 repeat: matches 2627..2750 of consensus"
32816..32979
/note="MIR repeat: matches 12..203 of consensus"
complement(33006..33450)
/note="MIR repeat: matches 91..259 of consensus"
complement(33058..33444)
/note="match: GSS: Em:A0261227"
33312..33506
/note="match: GSS: Em:A0040453"
34856..34915
/note="LIM4 repeat: matches 3473..3773 of consensus"
complement(34987..35547)
/note="match: GSS: Em:A0321598"
match: STR: Em:G55571"
complement(36468..36936)
/note="match: GSS: Em:BA1073"
complement(36654..37090)
/note="match: GSS: Em:A0772584"
37102..37591
/note="match: GSS: Em:A0489567"
complement(37581..38221)
/note="match: GSS: Em:A0418233"
37720..37772
repeat_region
```

misc_feature /note="12 repeat: matches 2657. .2710 of consensus"
complement(37778. .38215)
/note="match: GSS: Em:A0819773"
misc_feature 38218. .38732
/note="match: GSS: Em:A0599868"
repeat_region 38861. .39164
/note="A1usK repeat: matches 1. .305 of consensus"
repeat_region 39612. .39653
/note="21 copies 2 mer ac 90% conserved"
repeat_region 40854. .40909
/note="12 repeat: matches 2655. .2710 of consensus"
repeat_region 42169. .42270
/note="MER5A repeat: matches 1. .111 of consensus"
repeat_region 42411. .42548
/note="MER5B repeat: matches 21. .181 of consensus"
repeat_region 42780. .42938
/note="MER5B repeat: matches 17. .177 of consensus"
repeat_region 43229. .43326
/note="WIR repeat: matches 40. .148 of consensus"
repeat_region 45383. .45428
/note="MER5B repeat: matches 18. .62 of consensus"
misc_feature /note="match: STS: Em:G51995
match: GSS: Em:A0269059"
misc_feature /note="match: STS: Em:G36899"
complement(47537. .48011)
/note="match: STS: Em:G30049"
47972. .48451
/note="match: GSS: Em:A0336527"
complement(48518. .48985)
/note="match: GSS: Em:A0801834"
misc_feature 48597. .48803
repeat_region

Query Match 99.7%; Score 2814; DB 9; Length 80247;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2817; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 AAGCTTGTCTACGAGAGTCATATGCTTAAAGAAAACATACCAAAAGAGT 60
34271 AAGCTTGTCTACGAGAGTCATATGCTTAAAGAAAACATACCAAAAGAGT 34330
61 GCTTTTAAATGTTACATATATTTAGTAATTAATGAGAAATTTCCGCTTT 120
34331 GCTTTTAAATGTTACATATATTTAGTAATTAATGAGAAATTTCCGCTTT 34390
121 TGGGAAGAAACATCTAATGTTCTAGTTCTGCAAAATTTCAAAATCTTAA 180
34391 TGGGAAGAAACATCTAATGTTCTAGTTCTGCAAAATTTCAAAATCTTAA 34450
181 TGAGTCTGAATTTATAGCCATATTTTGTGTAAGAGAACATGAATTTCAATG 240
34451 TGAGTCTGAATTTATAGCCATATTTTGTGTAAGAGAACATGAATTTCAATG 34510
241 ACCCGTGAGAAATTTATCAATAAAGTTTTCGCAAAATATATTTTGTCTATAGCCAT 300
34511 ACCCGTGAGAAATTTATCAATAAAGTTTTCGCAAAATATATTTTGTCTATAGCCAT 34570
301 AAAGAGCATGAATGCGGAAGAAAAAACTGAAACACTTTTCTTAAAAAAAAA 360
34571 AAAGAGCATGAATGCGGAAGAAAAAACTGAAACACTTTTCTTAAAAAAAAA 34630
361 GCGTCACATCTTTTATAGCCAAATCTGTAATTAATCATGACATCAAGTTGTTCTTC 420
34631 GCGTCACATCTTTTATAGCCAAATCTGTAATTAATCATGACATCAAGTTGTTCTTC 34690
421 ATGAATATAGCCCTTTTATAGCAATTTGTAATTTTAAATACGATATCATCTTAAT 480
34691 ATGAATATAGCCCTTTTATAGCAATTTGTAATTTTAAATACGATATCATCTTAAT 34750
481 ACAAAGAAAAAATTTTAAAGCCCTTAGGCTATGTAATGAGAAAGTTTCCATGCA 540
34751 ACAAAGAAAAAATTTTAAAGCCCTTAGGCTATGTAATGAGAAAGTTTCCATGCA 34810

541 AATGAAACATTTGATTTATACATCTTGTAGATATGAGATGTAACAATTTATTA 600
34811 AATGAAACATTTGATTTATACATCTTGTAGATATGAGATGTAACAATTTATTA 34870
601 GTAACAAAGAAAACGTAGAAAATATTAATTAATTAATTAATTAATTAATTA 660
34871 GTAACAAAGAAAACGTAGAAAATATTAATTAATTAATTAATTAATTAATTA 34930
661 ATTTATAAATTTGCTGAGTGTAGTTTAAATATGTAATTAATTAATTAATTA 720
34931 ATTTATAAATTTGCTGAGTGTAGTTTAAATATGTAATTAATTAATTAATTA 34990
721 ATTTTATACAGATATTTATTAATGATTTGCAAACTGTAGAGTAATTTATAG 780
34991 ATTTTATACAGATATTTATTAATGATTTGCAAACTGTAGAGTAATTTATAG 35050
781 AAAAAAACAACATTAACGTTTAAACAGTTTAAACAGTTTAAACAGTTTAAAC 840
35051 AAAAAAACAACATTAACGTTTAAACAGTTTAAACAGTTTAAACAGTTTAAAC 35110
841 ACCAAGAGTATGTAAGTGTATGATATGCGAGTAATTAATGATATTAATCACTGGA 900
35111 ACCAAGAGTATGTAAGTGTATGATATGCGAGTAATTAATGATATTAATCACTGGA 35170
901 GATGTTCAACAAATCTGATGTTGTTACAAATTAATTAATTAATTAATTAATTA 960
35171 GATGTTCAACAAATCTGATGTTGTTACAAATTAATTAATTAATTAATTAATTA 35230
961 TGACAGCTAGTGTGTGCACTTTCTACAGCAAAATTAATTAATTAATTAATTA 1020
35231 TGACAGCTAGTGTGTGCACTTTCTACAGCAAAATTAATTAATTAATTAATTA 35290
1021 CATTGAAAAAAGCTTACTGTTTGAATATGCTGATTTTGAAGAGTTTGGG 1080
35291 CATTGAAAAAAGCTTACTGTTTGAATATGCTGATTTTGAAGAGTTTGGG 35350
1081 AAACCTCTTTTGAAGAGTGTGCTGCTTGAAGAAAGTCTGTTCTGCAATGAT 1140
35351 AAACCTCTTTTGAAGAGTGTGCTGCTTGAAGAAAGTCTGTTCTGCAATGAT 35410
1141 TGATTAACCTGAAGAAAGAGTGTGATTAATTAATTAATTAATTAATTAATTA 1200
35411 TGATTAACCTGAAGAAAGAGTGTGATTAATTAATTAATTAATTAATTAATTA 35470
1201 TTAATATATTCGCAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
35471 TTAATATATTCGCAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 35530
1261 TGTCTTCTCTATTTTGTAGAAAATTTATGACATCAATCAAGAAATTTCTGCCC 1320
35531 TGTCTTCTCTATTTTGTAGAAAATTTATGACATCAATCAAGAAATTTCTGCCC 35590
1321 ATTTTAAATGCTGATTAATCTGGGCTGTTTATGATGATGATGATGATGATGAT 1380
35591 ATTTTAAATGCTGATTAATCTGGGCTGTTTATGATGATGATGATGATGATGAT 35650
1381 ACCTACCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
35651 ACCTACCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 35710
1441 AAACCTCTTTTCAACCAAAAGATTAATTTATGAAATCTAGATCTGCTGGGCTGG 1500
35711 AAACCTCTTTTCAACCAAAAGATTAATTTATGAAATCTAGATCTGCTGGGCTGG 35770
1501 GCCATTTGGGAGACCTTTGACCTTAATCACTGCAAGCAATGTTCAACATCAAGATG 1560
35771 GCCATTTGGGAGACCTTTGACCTTAATCACTGCAAGCAATGTTCAACATCAAGATG 35830
1561 GCATGCAACAAGGAGCCGTTTCACTGTAACAACAAATCCCATGAGTTTGTATCATCAGG 1620
35831 GCATGCAACAAGGAGCCGTTTCACTGTAACAACAAATCCCATGAGTTTGTATCATCAGG 35890

1621 AGACATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 1680
 35991 AGCAGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 35950
 1681 GATTGGATGAAATTTGGATACCACTCAGACAGCTAGGACCCCTCCAGAAATG 1740
 35951 GATTGGATGAAATTTGGATACCACTCAGACAGCTAGGACCCCTCCAGAAATG 36010
 1741 GTTAACTGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 1800
 36011 GTTAACTGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36070
 1801 GAGGTGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 1860
 36071 GAGGTGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36130
 1861 GATTGGATGAAATTTGGATACCACTCAGACAGCTAGGACCCCTCCAGAAATG 1920
 36131 GATTGGATGAAATTTGGATACCACTCAGACAGCTAGGACCCCTCCAGAAATG 36190
 1921 AAGCATGAAATTCATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 1980
 36191 AAGCATGAAATTCATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36250
 1981 GATTGGATGAAATTTGGATACCACTCAGACAGCTAGGACCCCTCCAGAAATG 2040
 36251 GATTGGATGAAATTTGGATACCACTCAGACAGCTAGGACCCCTCCAGAAATG 36310
 2041 TTATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2100
 36311 TTATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36370
 2101 ACCATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2160
 36371 ACCATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36430
 2161 TTATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2220
 36431 TTATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36490
 2221 ATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2280
 36491 ATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36550
 2281 ATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2340
 36551 ATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36610
 2341 TAAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2400
 36611 TAAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36670
 2401 TTGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2460
 36671 TTGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36730
 2461 TCAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2520
 36731 TCAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36790
 2521 AAGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2580
 36791 AAGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36850
 2581 TGGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2640
 36851 TGGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36910
 2641 TTGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2700
 36911 TTGATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 36970
 2701 TTAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2760

36971 TTAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 37030
 2761 TTAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 2820
 37031 TTAATGAGTGGGATCTGACAAATTTACTCAGACAGCTAGGACCCCTCCAGAAATG 37090
 2821 TC 2822
 37091 TC 37092
 RESULT 2
 AB023021 3019 bp mRNA linear PRI 23-JUN-1999
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 FEATURES
 source
 gene
 CDS
 BASE COUNT
 ORIGIN
 Query Match 54.6%; Score 1542; DB 9; Length 3019;
 Best Local Similarity 100.0%; Pred. No. 2,46-261;
 Matches 1542; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1401 TCAGTCCATGGAATGACGAGCTGTGTGTAATAATGAAAACTCTTTCCACCAAAA 1460
 |||||||
 Db 407 TCAGTCCATGGAATGACGAGCTGTGTGTAATAATGAAAACTCTTTCCACCAAAA 466
 |||||||
 QY 1461 CTGATATTTTATGAACTACTATCTGTGTGGGTGTGGCATTTGGGCACACTTTG 1520
 |||||||
 Db 467 CTGATATTTTATGAACTACTATCTGTGTGGGTGTGGCATTTGGGCACACTTTG 526
 |||||||
 QY 1521 ACCTTACATCTGCGCAAGCAATGTTCACATCCCAAGGATGCCATCTCACAAAGGACCTT 1580
 |||||||
 Db 527 ACCTTACATCTGCGCAAGCAATGTTCACATCCCAAGGATGCCATCTCACAAAGGACCTT 586
 |||||||
 QY 1581 CACTGTACAACAAAATCCATGCAAGTCTGATCCATCACCAGACATCAGTTGGGATCTGA 1640
 |||||||
 Db 587 CACTGTACAACAAAATCCATGCAAGTCTGATCCATCACCAGACATCAGTTGGGATCTGA 646
 |||||||
 QY 1641 CAATTTACTCTAGCAAGAGTACGCGCACCCTCCAGAAATGGATTTGGATGAAATTTGGAA 1700
 |||||||
 Db 647 CAATTTACTCTAGCAAGAGTACGCGCACCCTCCAGAAATGGATTTGGATGAAATTTGGAA 706
 |||||||
 QY 1701 CACCAACTCACACTCCCAAAAAGAGTGAGCACTTGTTTAACTGTACTGTACTT 1760
 |||||||
 Db 707 CACCAACTCACACTCCCAAAAAGAGTGAGCACTTGTTTAACTGTACTGTACTT 766
 |||||||
 QY 1761 ACCGCCGTGATTCAGATATCCAAAGTCCCTTANGCTTCTTGACGTTAGCACAATCCCT 1820
 |||||||
 Db 767 ACCGCCGTGATTCAGATATCCAAAGTCCCTTANGCTTCTTGACGTTAGCACAATCCCT 826
 |||||||
 QY 1821 TCGTGTGTAAGTCCCAACCAAGAAATGTTGCTGCTGGGTGTGAGTAAGTACGTAAGC 1880
 |||||||
 Db 827 TCGTGTGTAAGTCCCAACCAAGAAATGTTGCTGCTGGGTGTGAGTAAGTACGTAAGC 886
 |||||||
 QY 1881 CTGAGCATCCAGAGTCAAGTATTTACATAGCTAAGCAAAAAGCATGGAATCCATACCT 1940
 |||||||
 Db 887 CTGAGCATCCAGAGTCAAGTATTTACATAGCTAAGCAAAAAGCATGGAATCCATACCT 946
 |||||||
 QY 1941 ACGGCGACGATTTGGAGAAATGTCATGATTAATAATTTGATCTCTACCATATCTGCTT 2000
 |||||||
 Db 947 ACGGCGACGATTTGGAGAAATGTCATGATTAATAATTTGATCTCTACCATATCTGCTT 1006
 |||||||
 QY 2001 GTTAAATTTATCTTCTCTTGAATAATTCATCCACAAAGATTCATCAGGAAAAACTAT 2060
 |||||||
 Db 1007 GTTAAATTTATCTTCTCTTGAATAATTCATCCACAAAGATTCATCAGGAAAAACTAT 1066
 |||||||
 QY 2061 ACAATGCTTTCTGCTGCTGCTGTACCTGTTCTGTGGACCATTTAGGAAAAACTATG 2120
 |||||||
 Db 1067 ACAATGCTTTCTGCTGCTGCTGTACCTGTTCTGTGGACCATTTAGGAAAAACTATG 1126
 |||||||
 QY 2121 AGAATTAATTCAGCAGATTCATTCATGCTGAGAAATTAATCTCCCAAGTAC 2180
 |||||||
 Db 1127 AGAATTAATTCAGCAGATTCATTCATGCTGAGAAATTAATCTCCCAAGTAC 1186
 |||||||
 QY 2181 TAGCAAGATTCGAAGGAAGTCGACAAAAACAATAAGTTARACCTTACTTTAACT 2240
 |||||||
 Db 1187 TAGCAAGATTCGAAGGAAGTCGACAAAAACAATAAGTTARACCTTACTTTAACT 1246
 |||||||
 QY 2241 GGAGGAAGATTTCACTGTAATCTCCAGATTTTGGGAATCCATGATGTTGGCTT 2300
 |||||||
 Db 1247 GGAGGAAGATTTCACTGTAATCTCCAGATTTTGGGAATCCATGATGTTGGCTT 1306
 |||||||
 QY 2301 GCGATCATGTGAAGGATCAAGAAATTAAGTCTGTTGGTAATTTAGAAAAATGCTTTT 2360
 |||||||
 Db 1307 GCGATCATGTGAAGGATCAAGAAATTAAGTCTGTTGGTAATTTAGAAAAATGCTTTT 1366
 |||||||
 QY 2361 GGAATTAATTTTCACTGTAATCTCCAGATTTTGGGAATCCATGATGTTGGCTT 2420
 |||||||
 Db 1367 GGAATTAATTTTCACTGTAATCTCCAGATTTTGGGAATCCATGATGTTGGCTT 1426
 |||||||
 QY 2421 GTATTGAGATTAAGAGATGCAACATACTTTTGTGCAATTTATTTTATCTAC 2480
 |||||||
 Db 1427 GTATTGAGATTAAGAGATGCAACATACTTTTGTGCAATTTATTTTATCTAC 1486
 |||||||

QY 2481 CCTCTAGGGTAACTGATATATTTGTGTGAGATTTTAAAAAGCTCAGCATGAGCAATC 2540
 |||||||
 Db 1487 CCTCTAGGGTAACTGATATATTTGTGTGAGATTTTAAAAAGCTCAGCATGAGCAATC 1546
 |||||||
 QY 2541 ATTCCATTTGGTTTAAATATCCCTGATATACCTAATTAATGACCTGAGAGTAATTT 2600
 |||||||
 Db 1547 ATTCCATTTGGTTTAAATATCCCTGATATATACCTAATTAATGACCTGAGAGTAATTT 1606
 |||||||
 QY 2601 ATTCCATTTATCATTTGTAACATTTGTTTTCACATTTTGTAGTTTCCATATGTA 2660
 |||||||
 Db 1607 ATTCCATTTATCATTTGTAACATTTGTTTTCACATTTTGTAGTTTCCATATGTA 1666
 |||||||
 QY 2661 AGCTTGTGTTGATTAATTTGTTTCCACACTGATCAGCTGTTAATCTAATTTGGGAATGA 2720
 |||||||
 Db 1667 AGCTTGTGTTGATTAATTTGTTTCCACACTGATCAGCTGTTAATCTAATTTGGGAATGA 1726
 |||||||
 QY 2721 AGATGCAATCTTAAAGTANGAAAAATTTTCCACTAGTATTAACAATGCTAGTCCAACT 2780
 |||||||
 Db 1727 AGATGCAATCTTAAAGTANGAAAAATTTTCCACTAGTATTAACAATGCTAGTCCAACT 1786
 |||||||
 QY 2781 TTGCATATCTATAACAGAGAGAAACATGTTGCATTTGAATTC 2822
 |||||||
 Db 1787 TTGCATATCTATAACAGAGAGAAACATGTTGCATTTGAATTC 1828
 |||||||

RESULT 3
 HSA238701
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

HSA238701 2501 bp mRNA linear PRI 21-OCT-2000
 Homo sapiens mRNA for alpha-3-fucosyltransferase.
 AJ238701.1 GI:4741566
 alpha-3-fucosyltransferase; FUT9 gene.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2501)
 Caillieu-Thomas,B., Coullin,P., Candelier,J.J., Balazsino,L.,
 Mennesson,B., Oriol,R. and Mollicone,R.
 FUT4 and FUT9 genes are expressed early in human embryogenesis
 Glycobiology 10 (8), 789-802 (2000)
 20386991

JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

2 (bases 1 to 2501)
 Mollicone,R.
 Direct Submission
 Submitted (29-Apr-1999) Mollicone R., Glycobiology, INSERM U504, 16
 Av. Paul Vaillant-Couturier, Villejuif, 94807 Cedex, FRANCE
 related sequence: AB015426.
 location/Qualifiers
 1..2501
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /rname_type="whole embryo"
 /dev_stage="40 to 60 days"
 105..2466
 /gene="FUT9"
 105..1184
 /gene="FUT9"
 /codon_start=1
 /product="alpha-3-fucosyltransferase"
 /protein_id="CAB41890.1"
 /db_xref="GI:4741567"
 /db_xref="SPTREMBL:O9Y231"
 /translation="MTSSKIKTLVWVFGQTFEDTSCQAMFNIOGCHLTDRLSYNK
 SSVLAKNFESTKDYENETILVWVFGQTFEDTSCQAMFNIOGCHLTDRLSYNK
 SHAVLIHRDLISWDLTNLPQOARPEFOKIMWNLESPTHPQKSGIEHLFNTLYR
 DSDIOVPGFLTVSNPEFVPSKEKLVCWVSNMNEHARVYVYNLSKSEIHTY
 GOARGEYVNDKMLPTTISACKFVLSFENSIRKDYITKRLYNAFLAGSVPVLAGSREN
 YENYIPADSPFIHVEDYNSPSELAKYIKVDKNNKLILYSIFWRRDFTVNLPRFESHA
 CLADHVRHDEIYSVGNLEKMPNN"
 2192
 /gene="FUT9"

polyA_site

/gene="FUT9"

```

/evidence=experimental
polyA_site
2235
/genes="PUT9"
/evidence=experimental
polyA_site
2310
/genes="PUT9"
/evidence=experimental
polyA_site
2466
/genes="PUT9"
/evidence=experimental
BASE COUNT      818 a      436 c      422 g      825 t
ORIGIN

```

```

Query Match      54.5%; Score 1537.2; DB 9; Length 2501;
Best Local Similarity 99.8%; Pred. No. 1.7e-260;
Matches 1539; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1281 GAAAAATTATGACATCAACATCCAAAGAAATCTTGGCCCATTTTAAATTTGTCGATTA 1340
|||||
Db 97 GAAAAATTATGACATCAACATCCAAAGAAATCTTGGCCCATTTTAAATTTGTCGATTA 156
OY 1341 TCGTGGGCTTTTCATGGCAGTCTCTTCATTTACATCAACCTAACCAACGCTGAGTCT 1400
|||||
Db 157 TCGTGGGCTTTTCATGGCAGTCTCTCTTCATTTACATCAACCTAACCAACGCTGAGTCT 216
OY 1401 TCAGTCCAAATGGAATCAGCCAGCTCTGTGCTGAAAATGAAAATCTTTTCCACANAA 1460
|||||
Db 217 TCAGTCCAAATGGAATCAGCCAGCTCTGTGCTGAAAATGAAAATCTTTTCCACANAA 276
OY 1461 CTGATTTATTTAAATGAATCTACTATCTGTGTGTGTGTGTCATTTGGGAGACCTTTG 1520
|||||
Db 277 CTGATTTATTTAAATGAATCTACTATCTGTGTGTGTGTGTCATTTGGGAGACCTTTG 336
OY 1521 ACCCTACATCCGCGCAAGCAATGTGTCAACATCCCAAGATGCCATCTCAACGAGCGTT 1580
|||||
Db 337 ACCCTACATCCGCGCAAGCAATGTGTCAACATCCCAAGATGCCATCTCAACGAGCGTT 396
OY 1581 CACTGTACAAACAATCCCATGAGTCTGATTCATCACGAGACATCAGTTGGATCTGA 1640
|||||
Db 397 CACTGTACAAACAATCCCATGAGTCTGATTCATCACGAGACATCAGTTGGATCTGA 456
OY 1641 CAAATTTACCTCAGCAAGCTGAGCCACCTTCCAGAAATGATTTGGATTTGGAAAT 1700
|||||
Db 457 CTAACTTACCTCAGCAAGCTGAGCCACCTTCCAGAAATGATTTGGATTTGGAAAT 516
OY 1701 CACCAACTCACACTCCCAAAAAGATGGCAATGAGCACTTGTTTAACTGACTGACTT 1760
|||||
Db 517 CACCAACTCACACTCCCAAAAAGATGGCAATGAGCACTTGTTTAACTGACTGACTT 576
OY 1761 ACCGCGGTGATTCAGATATCCAAAGTGCCTTATGGCTTCTTGACGGTAAAGCAAAATCCCT 1820
|||||
Db 577 ACCGCGGTGATTCAGATATCCAAAGTGCCTTATGGCTTCTTGACGGTAAAGCAAAATCCCT 636
OY 1821 TCGTGTGTAAGTGCAGCAAGAAAGAAATGGTGTGCGGGTGTGTGTAACGTAACGAAAC 1880
|||||
Db 637 TCGTGTGTAAGTGCAGCAAGAAAGAAATGGTGTGCGGGTGTGTGTAACGTAACGAAAC 656
OY 1881 CTGAGCATGCGAGACTCAAGTATTCATGAGCTAAAGCAAAAGCAATTTGAATTCATCT 1940
|||||
Db 697 CTGAGCATGCGAGACTCAAGTATTCATGAGCTAAAGCAAAAGCAATTTGAATTCATCTCT 756
OY 1941 ACGGCAAGCAATTTGAGAAATATGTCATGATTTAAATTTGATTTGCTTCCATATCTGCTT 2000
|||||
Db 757 ACGGCAAGCAATTTGAGAAATATGTCATGATTTAAATTTGATTTGCTTCCATATCTGCTT 816
OY 2001 GTAATTTTATCTTCTTCTTGAATAATCAATCCAAAGATATACATCCAGGAAAGATAT 2060
|||||
Db 817 GTAATTTTATCTTCTTCTTGAATAATCAATCCAAAGATATACATCCAGGAAAGATAT 876
OY 2061 ACAATGCTTTTCTGCTGCTCTGTACCTGTTGTTCTGAGACATCTAGGAAAAATATG 2120
|||||
Db 877 ACAATGCTTTTCTGCTGCTCTGTACCTGTTGTTCTGAGACATCTAGGAAAAATATG 936

```

```

OY 2121 AGAATTATATTCAGACATTCATTCATTCATGATGGAAGATTATTAATCTCCAGTAC 2180
|||||
Db 937 AGAATTATATTCAGACATTCATTCATTCATTCATGATGGAAGATTATTAATCTCCAGTAC 996
OY 2181 TAGCAAGATATTCAGAGAGTGCAGCAAAAGAAATTAATTAATTAATTAATTAATTA 2240
|||||
Db 997 TAGCAAGATATTCAGAGAGTGCAGCAAAAGAAATTAATTAATTAATTAATTAATTA 1056
OY 2241 GGAGAGAGATTTGACGTGAATCTTCAACGATTTTGGAAATCAATGCAATGTTGGCTT 2300
|||||
Db 1057 GGAGAGAGATTTGACGTGAATCTTCAACGATTTTGGAAATCAATGCAATGTTGGCTT 1116
OY 2301 GCGATCATGTGMAAAGCATCGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2360
|||||
Db 1117 GCGATCATGTGMAAAGCATCGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1176
OY 2361 GGAATTTAAATTTTTCATCTGACACTTGCACATTTGATTAATTTGATGATATATCCAA 2420
|||||
Db 1177 GGAATTTAAATTTTTCATCTGACACTTGCACATTTGATTAATTTGATGATATATCCAA 1236
OY 2421 GTATTGAGATTAAGAGATGCAACATCTACTTTTGTGTCACATTTTATTTTATCAC 2480
|||||
Db 1237 GTATTGAGATTAAGAGATGCAACATCTACTTTTGTGTCACATTTTATTTTATCAC 1296
OY 2481 CCTCTTAGGATTAAGAGATGCAACATCTACTTTTGTGTCACATTTTATTTTATCAC 2540
|||||
Db 1297 CCTCTTAGGATTAAGAGATGCAACATCTACTTTTGTGTCACATTTTATTTTATCAC 1356
OY 2541 ATTTCATTCGCTTTTAAATTAATTCCTGATATACCTAATTTATGTCACATGAGATTAAT 2600
|||||
Db 1357 ATTTCATTCGCTTTTAAATTAATTCCTGATATACCTAATTTATGTCACATGAGATTAAT 1416
OY 2601 ATTTCATTCATTTTAAATTAATTCCTGATATACCTAATTTATGTCACATGAGATTAAT 2660
|||||
Db 1417 ATTTCATTCATTTTAAATTAATTCCTGATATACCTAATTTATGTCACATGAGATTAAT 1476
OY 2661 AGCTGTGTTGATTAATTTGTCACATGATGATGATTTTATTAATTTGGAATTA 2720
|||||
Db 1477 AGCTGTGTTGATTAATTTGTCACATGATGATGATTTTATTAATTTGGAATTA 1536
OY 2721 AGATGACATCTTAAAGATGAATAATTTTCACTAAGATTAATTAATTTGGAATTA 2780
|||||
Db 1537 AGATGACATCTTAAAGATGAATAATTTTCACTAAGATTAATTAATTTGGAATTA 1596
OY 2781 TTGCATTAATTAAGAGAGAAACATGTTGCAATTTGAATTC 2822
|||||
Db 1597 TTGCATTAATTAAGAGAGAAACATGTTGCAATTTGAATTC 1638

```

```

RESULT 4
AF230460 1705 bp mRNA linear ROD 12-JUL-2000
LOCUS
DEFINITION
Cricketus gryseus alpha(1.3) fucosyltransferase type IX (Fut9)
ACCESSION
AF230460
VERSION
AF230460.1 GI:9049663
KEYWORDS
SOURCE
ORGANISM
Chinese hamster.
Cricketus gryseus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 1705)
REFERENCE
1 Patnaik, S.K., Zhang, A., Shi, S. and Stanley, P.
Alphal(1,3)fucosyltransferase expressed by the gain-of-function
Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30
JOURNAL
MEDLINE
20186933
PUBMED
10700388
2 (bases 1 to 1705)
REFERENCE
1 Patnaik, S.K., Shi, S. and Stanley, P.
Direct Submission
JOURNAL
Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of

```

Medicine, Chanin 516, 1300 Morris Park Avenue, New York, NY 10461, USA

FEATURES

Location/Qualifiers

1. .1705

/organism="Cricetulus griseus"

/db_xref="taxon:10029"

1. .1705

/gene="Fut9"

/note="cgFut9"

62. .1141

/gene="Fut9"

/note="Fuc-TIX"

/codon_start=1

/product="alpha(1,3)fucosyltransferase type IX"

/protein_id="AAF82412.1"

/db_xref="GI:3049664"

/translation="METSRSKGLRPLIVCVLIACFMACLLIYIKPTNSWSPWESA
SSVLKMKRPFSTKDYENETIILVYVMPFGOTFDTSCQAFNIDGCHLTDRSLYRK
SHAVLIHRDISMDLTNLPQARPPQKIMNNLESPTHTPOKSGIEHLNLTLYRK
DSDIQVPGFLTVSTNPPEVPSKELVCMWVSNMNEHARVKYKXLEHLSKSIETLY
GOAFGEYVNERNLPTISTCKFYLSEFNSIKHDYTEKLYNAFLAGSVPVVLGPREN
YENIIPADSFHVEDYNSPELAKYLFVDKNNKLYLSYFWMRKDFVNLPRFMSHA
CLACDHYVRHDEYSVGNLEKMFWM"

CDS

BASE COUNT 515 a 336 c 320 g 534 t

ORIGIN

Query Match 39.1%; Score 1103.2; DB 10; Length 1705;
Best Local Similarity 85.4%; Pred. No. 3e-184;
Matches 1301; Conservative 0; Mismatches 213; Indels 10; Gaps 6;

1282 AAAATTATGACATCAACATCCAAAGAAATCTCGCCCATTTTAATTGCTGATTAAT 1341
|||||
55 AAAAATTATGACATCAACATCCAAAGAAATCTCGCCCATTTTAATTGCTGATTAAT 114
|||||
1342 CCGGGCGTTCATGCGCATGCTCTCTCATTTACATCAACCACTACCAAGCTGATCTT 1401
|||||
115 CCGGCGTTCATGCGCATGCTCTCTCATTTACATCAACCACTACCAAGCTGATCTT 174
|||||
1402 CAGTCCAAATGGAATGACGACGCTGTGTGAAATGAAATCTTTTCCACCAAAAC 1461
|||||
175 CAGTCCAAATGGAATGACGACGCTGTGTGAAATGAAATCTTTTCCACCAAAAC 234
|||||
1462 TGAATATTTTAAATGAAATCTATCTGTGTGTGGTGTGGGATTTGGGACACCTTTGA 1521
|||||
235 TGAATATTTTAAATGAAATCTATCTGTGTGTGGTGTGGGATTTGGGACACCTTTGA 294
|||||
1522 CCTTCATCCTCCGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAGGACCGCTTC 1581
|||||
295 CCTTCATCCTCCGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAGGACCGCTTC 354
|||||
1582 ACTGTACAACAATCCATCCATGCTGTGATCCATACCGAGACATCAGTTGGATCTGCAC 1641
|||||
355 ACTGTACAACAATCCATCCATGCTGTGATCCATACCGAGACATCAGTTGGATCTGCAC 414
|||||
1642 AAATTTACCTGACAGAGCTAGGCCACCTTCAGAAATGATTTGGATTTGGATTC 1701
|||||
415 TAACTTACCTGACAGAGCTAGGCCACCTTCAGAAATGATTTGGATTTGGATTCAGTTC 474
|||||
1702 ACCAATCCACAGTCCGCAAAAGAGGATTTGGATTTGGATTTGGATTTGGATTTGGATTT 1761
|||||
475 TCCCACTCATACCCCAAAAGAGGATTTGGATTTGGATTTGGATTTGGATTTGGATTT 534
|||||
1762 CCGCGGTGATTCAGATATTCAGATGCTTATGCTTTCAGAGTACAGCAAAATCCCTT 1821
|||||
535 TCCCGGTGATTCAGATATTCAGATGCTTATGCTTTCAGAGTACAGCAAAATCCCTT 594
|||||
1822 CGTGTGTAAGTGCAGCAAGCAAGAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1881
|||||
595 TGTGTGTAAGTGCAGCAAGCAAGAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 654
|||||
1882 TGAGATGCCAGATCAAGTATTAAGATGAGTAAAGCAAAATGGAATTCATACCTTA 1941
|||||
655 GGAGATGCCAGATCAAGTATTAAGATGAGTAAAGCAAGATTTGATTTGATTTGATTTGATTT 714
|||||

1942 CCGGCAAGCATTTGGAGAAATATGCAATGATAAAATTTGATTCCTACCATATCTCTTG 2001
|||||
715 TGGGCAAGCCTTTGGAGAAATAGTGAATGAAAAAATTCATATCCACAGATATCTACTTG 774
|||||
2002 TAAATTTTATCTTCTCTTTGAAAATTCATTCACAGAGATTAACATCAGGAAAGCTATA 2061
|||||
775 TAAATTTTATCTTCTTCTTTGAAAATTCATTCACAGAGATTAACATCAGGAAAGCTATA 834
|||||
2062 CAATGCTTTTGGCGGCTGTACCTGT 2121
|||||
835 CAATGCTTTTGGCGGCTGTACCTGT 894
|||||
2122 GAATTAATATTCACAGAGATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2181
|||||
895 GAATTAATATTCACAGAGATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 954
|||||
2182 ACCAAGATATTCAGAGAGATTCAGCAAAACAAATTAATACCTTAGTTACTTAACTG 2241
|||||
955 GGCAGAAATATCTTAAAGGAATGACAAAACAAATTAATACCTTAGTTACTTAACTG 1014
|||||
2242 GAGGAAGATTTACCTGTAATCTTCCAGATTTGGAGATTCACATGCAATGCTTGGCTTG 2301
|||||
1015 GAGGAAGATTTACCTGTAATCTTCCAGATTTGGAGATTCACATGCAATGCTTGGCTTG 1074
|||||
2302 CGATCATGTGAAAAGGACATCAAGAAATATTAAGTCTGTGTGTAATTTAGAGAAATGCTTTTG 2361
|||||
1075 TGAATCATGTAAGAAAGGACATCAAGAAATATTAAGTCTGTGTGTAATTTAGAGAAATGCTTTTG 1134
|||||
2362 GAATTAATATTTTTCATCTGACACTTGAATTAATTTTGAATGATATCATTCACAG 2421
|||||
1135 GAATTAATATTTTTCATCTGACACTTGAATTAATTTTGAATGATATCATTCACAG 1194
|||||
2422 TATGAGATTAAGAGATTCACATCACTACTTTTGTGTCAATTTATTTTATTCACC 2481
|||||
1195 TGTGAGATTAAGAGATTCACATCACTACTTTTGTGTCAATTTATTTTATTCACC 1251
|||||
2482 CTCCTAGAGATTAAGAGATTCACATCACTACTTTTGTGTCAATTTATTTTATTCACC 2540
|||||
1252 CTCCTAGAGATTAAGAGATTCACATCACTACTTTTGTGTCAATTTATTTTATTCACC 1311
|||||
2541 ATTCATTCGCTTTTAAATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 2600
|||||
1312 ATTCATTCGCTTTTAAATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 1369
|||||
2601 ATTCATTCGCTTTTAAATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 2660
|||||
1370 ATTCATTCGCTTTTAAATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 1428
|||||
2661 ACCTTGAGTGTGATTAATTTGATGAGAGATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 2720
|||||
1429 ACCTTGAGTGTGATTAATTTGATGAGAGATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 1485
|||||
2721 AGATGACATCTTAAAGATTAAGTAAATTTTCACTAAGATTAATTCCTGATATACCTTAATTT 2780
|||||
1486 AAATGAGTGTGATTAATTTGATGAGAGATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 1545
|||||
2781 TTGCTACTATTAAGAGAGATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 2804
|||||
1546 TTGCTACTATTAAGAGAGATTAATTCCTGATATACCTTAATTTGATGAGAGATTAATTT 1569
|||||

RESULT 5
AB015426
LOCUS
DEFINITION
MUS musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.
ACCESSION
AB015426
VERSION
AB015426.1 GI:3702718
KEYWORDS
Fut9; alpha1,3-fucosyltransferase IX.
SOURCE
MUS musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
Mammalia: Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.							
1 (sites)	Kudo, T., Ikehara, Y., Togayachi, A., Kaneko, M., Hiraiga, T., Sasaki, K.	Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mFuc-1X, that synthesizes the Lewis x (CD15), epitope in brain and kidney	U. Biol. Chem. 273 (41), 26729-26738 (1998)				
2 (bases 1 to 2139)	Kudo, T. and Narimatsu, H.		98434388				
Direct Submission							
Submitted (09-JUN-1998)	Takashi Kudo, Institute of Life Science, Soka University, Division of Cell Biology, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan (E-mail: kudo@t.soka.ac.jp, Tel: +81-426-91-2495, Fax: +81-426-91-9315)						
Location/Qualifiers							
1. 2139							
/organism="Mus musculus"							
/db_xref="taxon:10090"							
115..1194							
/gene="Fut9"							
115..1194							
/gene="Fut9"							
/codon_start=1							
/product="alpha1,3-fucosyltransferase IX"							
/protein_id="BA03522.1"							
/db_xref="GI:3702719"							
/translation="MTNSKSGILRPILVCIIVCIILGCFNACILIIYIKPTNSWFSPEMSASVYLKKNKFEKTDYFNETTILVWVWPQGFDTLSCQALFNIOGCHLTRSLYNNSHVILHHNDDISMD/TLNLPQDARPPKQKIMWNLSPEHPQSGIEHLNFTLIRYDSIOVPYGLFVSTNPEVEPSKELICVWVSNMNEARVRYNELSIEIHTYRGGYFVNDKMLPTISTCKFLSEFENSIHDIYTEKLYNAELASVYVIGPREREKVEYFIADSFIFIVEDPNSPRLAKYLEVDKNNKLYLSYFMRDFTVNLPPFMESHACLCADHVRRHQRKQKSVGLKQFWFN"							
2139							
/note="31 a nucleotide"							
BASE COUNT	684 a	395 c	368 g 692 t				
ORIGIN							
polyA_site							
Query Match	38.7%; Score 1092.2; DB 10; Length 2139;						
Best Local Similarity	85.8%; Pval: No. 2.3e-182;						
Matches 1309; Conservative	0; Mismatches 198; Indels 18; Gaps 8;						
OY 1281	GAATAATTGACATCAACATCCAAAGAAATTCGCGCCATTTTAATGTCTGCATTA	1340					
Db 107	GAATAATTGACATCAACATCCAAAGAAATTCGCGCCATTTTAATGTCTGCATCA	166					
OY 1341	TCCTGGCGTTCATGGCATGTCTCTCATTTTACATCAAACTACCAACAGCTGATCT	1400					
Db 167	TCCTGGCGTTCATGGCATGTCTCTCATTTTATCAAGCCCAACACAGCTGGGCT	226					
OY 1401	TCAGTCAATGGAATCAGCAGCTCTGTGCTGAAAAATGAAAAAATCTTTCACACAAA	1460					
Db 227	TCAGTCAATGGAATCAGCAGCTCTGTGCTGAAAAATGAAAAAATCTTTCACAAAAA	286					
OY 1461	CTGATATTATTAAGAACTACATCTTCGCTGGGCTGGCCATTTGGGCGAGACCTTG	1520					
Db 287	CTGATATTATTAAGAACTACATCTTCGCTGGGCTGGCCATTTGGGCGAGACCTTG	346					
OY 1521	ACCTTACATCTGCCAAGCAATTTCAACATCCAAAGATCCATCTCACACGAGCGTT	1580					
Db 347	ACCTTACATCTGCCAAGCAATTTCAATTCAAAGGCTGCCATCTCACAAAGACGCT	406					
OY 1581	CAGTGCACAAATCCCATGCAAGTTCGATCCATCAACGAGACATCAATGGGATCTGA	1640					
Db 407	CAGTGCACAAATCCCATGCGTCTGATTCACATTAAGACATCACTGGATCGA	466					
OY 1641	CAAAATTTACCTACAGCTAGCGCACCGCTTCGAAAGAAATGATTTGGATTTGGAT	1700					
Db 467	CTAACTTACCTACAGCGCAGCGCACCGCTTTCAAGAAATGATTTGGATTTAAGAT	526					
OY 1701	CACCAATCACTACCTCCCAAAAGATGGCATTTGACACTTTGTTAACTGACTCTGACTT	1760					

D	527	CACCCTACACACCCCCAAAGATGTCATTTGGAACCTGTTCACCTGACTTACTACT	586
Q	1761	ACCCGCGTGAATGACATATCCAGTGGCTTATGCTCTTGACGGTAAGCAAAATCCCT	1820
D	587	ATCGCGCGATGATGATATATCCAAATGGCTTAAGGCTCTTGACGGGTGAGCAAAATCCCT	646
Q	1821	TGCTGTGGAAGTCCCAAGCAAAAGAAATTTGGTGGCTGGGTGTGTGATCTGCAAC	1880
D	647	TTGTGTGGAAGTCCCAAGCAAGAGAGAGTGGTGGCTGGCTGGCTGTGGATACGTGAAC	706
Q	1881	CTGACATGCGCACAGTCAAGATTTACAAAGACCTTAAGCAAAACATATAAATCCATACCT	1940
D	707	CTGACATGCGCACAGGTCAGATTTACACAGACCTCAGCAAGATATGTAATCCACACCT	766
Q	1941	ACGGGCAAGCAATTTGGAGATATCTCAATGATAAAAATTTGATATCCCGACATATGCTT	2000
D	767	ATGGCAAGCAATTCGSAAGATACGTGACAGATAAAAATCTGATCCACATATCTACTT	826
Q	2001	GTAAATTTTATCTCTTCTTGAATTCATCTCACAGATTTACATCACGSAAGATAT	2060
D	827	GTAAATTTTATCTCTTCTTGAATTCATCTCACAGATTTACATCACAGATATCTCT	886
Q	2061	ACATGCTTTTCTGCTGGCTCTGTACTGCTGTGTCTGGAGCCATCTAGGSAAAATATG	2120
D	887	ACAAATGCAATTTTGGCTGTGCTGATACCTGCTGTGTCTGGATCCATCTAGGSAAAATATG	946
Q	2121	AGAAATTTATTCACAGCAGATTTCAATCTCAATGATGGAAGATTTATACCTCCACAGTAC	2180
D	947	AGAAATTTATTCACAGCTGTATCTCAATCTCAATGATGGAAGATTTATACCTCCACAGTAC	1006
Q	2181	TAGCAAGATATCTGAAGAGAGTCAACAAAAAATAAGTATATCTGTGTACTTACT	2240
D	1007	TAGCAAAATATCTGAAGAGAGTCAACAAAAAATAAGTATATCTGTGTACTTACT	1066
Q	2241	GGAGAGAGATTTCACTGTAAATCTCCAGATTTTGGGAATACACAGATTTGGCT	2300
D	1067	GGAGAGAGATTTTACTGTAAACCTACACGGTTTTTGGGAATACAGATATGGCTGGAC	1128
Q	2301	GGCATCATGTGAAAGGCATCAGAAATATTAAGTCTGTGTGCTAATTTAGAGATGGTTT	2360
D	1127	GGACACATGTAAAGGCATCAGAAATATTAAGTCTGTGTGCTAATTTAGAGATGGTTT	1186
Q	2361	GGAAATTTAAATTTTTCATCACTTSCAACCTGTAAATATTTTGGATGATATCATCAA	2420
D	1187	GGAAATTTAAAGTCTCATCATCTGTGCACA--GTGAAAAATTTAAAGACATATTTCAA	1242
Q	2421	GTATTGAGAGATTAAGAGATGCAACATCTCTTTGTGTGCATTTATTTTATAC	2480
D	1243	GTTTTCAGAGATTAAGAGA----AACGCGTCAATTTGGGCAACGTTAATTTTCCTGC	1297
Q	2481	CCGCTCTAGGAGTACGTATATTTTGGGAGATTTTAAAGCTCAGACATGCAAC	2540
D	1298	CCGCTCTAGGAGCACTGTATATTTTGGTGA-ATTTTAAAGATCAAAATTTGCAATC	1356
Q	2541	ATTCATTTCCGTTTAAATATCTGTATATATCTATATATGATGCACTGGAGATATTT	2600
D	1357	ATTCATTTTGGTTTAAATATCTGTATATATGTAATGGAACGACACCTGGA-AATATTTT	1415
Q	2601	ATTCCTCATATCATTTGTAAACATTTGCTTTTTCACATTTTGTAGTGTGCATATATGA	2660
D	1416	ATTCCTCATCTCATTTAATAAATCTTTTTTTACA--TTTATATGCTGACTGTAAGTA	1477
Q	2661	ACCTTGTGGTATATATGTTTCCACACATGATACGCTTATATCTTTTGGAAATGA	2720
D	1473	AATTTATGAT--TACTGTTTCTACATCAATTCAGATCTTATCTATTTGGAGATA	1528
Q	2721	AATGTCACATCTTAAAGATGA-AAAATTTCAAGATATATCAATGCTATGCTCAAC	2779
D	1530	AATATCATATCTTAAATATGAGGACCTTTTGGCCAAATATTTAATAGATATGATCTGAG	1588
Q	2780	TTTGCACTATTAACAGAGGAAGA 2804	

Db 1590 CTTGTAATGCTACAAAGAGAGA 1614

RESULT 6
AB049819 2156 bp mRNA linear ROD 18-JAN-2002
LOCUS Rattus norvegicus FUT9 mRNA for alpha1,3-fucosyltransferase IX,
DEFINITION complete cds.
ACCESSION AB049819
VERSION AB049819.1 GI:13591588
KEYWORDS Rattus norvegicus CDNA to mRNA.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (sites)
AUTHORS Shimoda,Y., Tajima,Y., Osanai,T., Katsume,A., Kohara,M., Kudo,T.,
Narimatsu,H., Takashima,N., Ishii,Y., Nakamura,S., Osuni,N. and
Sanai,Y.
TITLE Pax6 Controls the Expression of Lewis x Epitope in the Embryonic
Forebrain by Regulating alpha 1,3-Fucosyltransferase IX Expression
JOURNAL J. Biol. Chem. 277 (3), 2033-2039 (2002)
PUBMED 11675393
REFERENCE 2 (bases 1 to 2156)
AUTHORS Sanai,Y.
TITLE Direct Submission
REFERENCE Submitted (11-OCT-2000) Yutaka Sanai, Tokyo Metropolitan Institute
of Medical Science, Department of Biochemical Cell Research;
Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113-8613, Japan
(E-mail:sanai@ishoken.or.jp, Tel:81-3-3823-2101(ex.5233),
Fax:81-3-3828-6663)
FEATURES
source location/Qualifiers
1. 2156
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
346..1425
/gene="FUT9"
346..1425
/gene="FUT9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="BAB40953.1"
/db_xref="GI:13591589"
CDS
translation="MSTSGGILIRPLIVCIILGCFVACILYIKPNSWVSPMESA
SSVLKMKNFPSKRTDYFNETTLLVWVPGQFDFLSCAMNIOGCHLTITRSLYRK
SHAVLIHHDIDMDLNLNLPQARPPROKIMNLSPPTPQSGIEHLFNTITVRR
DSDIQYPIGFLVYSTSPYFEPYPSKELKLYVSNMPEHAKYITNELSKSEIHTY
GQAFGEYVNDKMLIPTISTCKFYLSKFIYSENIHKDYITERLYNAFLAGSVPLGPSKEN
YENYIPADSFIVHEDFNSESLAKYLIKEVDKNNKLYLSYNNRKDETVNLPRWESHA
CLACDHVKHOKESYGNLEKWFVN"
BASE COUNT 605 a 470 c 437 g 644 t
ORIGIN

Query Match 38.5%; Score 1085.4; DB 10; Length 2156;
Best Local Similarity 84.3%; Pred. No. 3,9e-181;
Matches 1285; Conservative 0; Mismatches 226; Indels 14; Gaps 5;

QY 1281 GAAAAATTATGACATCAACATCCAAAGAGATTCCTGCCCATTTTAAATGTCGCACTTA 1340
DB 338 GAAAAATTATGACATCAACATCCAAAGAGATTCCTGCCCATTTTAAATGTCGCACTTA 397
QY 1341 TCGTGGCGTGTTCATGCGATGCTCTCTCATTTACATCAACATCCAAAGAGATTCCTGCCCATTTTAAATGTCGCACTTA 1400
DB 398 TCGTGGCGTGTTCATGCGATGCTCTCTCATTTACATCAACATCCAAAGAGATTCCTGCCCATTTTAAATGTCGCACTTA 457
QY 1401 TCGATCCAAATGGAATGACGACGCTCTGCTGAAATGAAAACTTTTCCACCAAAA 1460
DB 458 TCGATCCAAATGGAATGACGACGCTCTGCTGAAATGAAAACTTTTCCACCAAAA 517
QY 1461 CTGATTTATTTTAAAGAACTACTATTTCTGCTGGGTGGGCGCATTTGGGCGACGCTTTG 1520
DB 518 CTGATTTATTTTAAAGAACTACTATTTCTGCTGGGTGGGCGCATTTGGGCGACGCTTTG 577

QY 1521 ACCTTACATCCTGCGAAGCAATGTTCAACATCCAAAGAGATCCATCTCACACGACCGTT 1580
DB 578 ACCTTACATCCTGCGAAGCAATGTTCAACATCCAAAGAGATCCATCTCACACGACCGTT 637
QY 1581 CACTGTCAACAAATCCCATGACGCTTGTATCCATCCAGACATCAGTGGATGTA 1640
DB 638 CGTGTCAACAAATCCCATGACGCTTGTATCCATCCAGACATCAGTGGATGTA 697
QY 1641 CAATTTACCTGCAAGCAAGTACGACGCTTCCGAAATGGAATTTGGATGTAATTTGGAAT 1700
DB 698 CTACTTACCTGCAAGCAAGTACGACGCTTCCGAAATGGAATTTGGATGTAATTTGGAAT 757
QY 1701 CACCACTCACTGCTCCCAAAAGAGTGGATTTGAGCACTTTGTAACCTACTCTGACT 1760
DB 758 CACCACTCACTGCTCCCAAAAGAGTGGATTTGAGCACTTTGTAACCTACTCTGACT 817
QY 1761 ACCGCGCTGATTCAGATATTCAGATGCTTATGCTCTTGCAGGTAGCAAAATCCCT 1820
DB 818 ATGCGCGTGAATTCAGATATTCAGATGCTTATGCTCTTGCAGGTAGCAAAATCCCT 877
QY 1821 TCGGTGTTGAAGTCCCAAGCAAGAAATTTGGTGTCTGGGTTTGGATGTAATCCGAAAC 1880
DB 878 TCGGTGTTGAAGTCCCAAGCAAGAAATTTGGTGTCTGGGTTTGGATGTAATCCGAAAC 937
QY 1881 CTGAGCATGCGCAGATCAAGTATTCAGATGCTTAAAGCAAAAGCATTTGAATCCATACT 1940
DB 938 CCGAGCATGCGCAGATCAAGTATTCAGATGCTTAAAGCAAAAGCATTTGAATCCATACT 997
QY 1941 ACGGCGAAGCATTTGGAGATATGTCATATATAAATTTGATCTCTACATATGCTT 2000
DB 998 ATGCGCGTGAATTCAGATATTCAGATGCTTAAATGATCCATCCATCTACTTACTT 1057
QY 2001 GTAAATTTATCTTCTCTGTAAGAAATTCATCCAAAGATTTACATCCGAAAGATAT 2060
DB 1058 GTAAATTTATCTTCTCTGTAAGAAATTCATCCAAAGATTTACATCCGAAAGATAT 1117
QY 2061 ACAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120
DB 1118 ACAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
QY 2121 AGAATTTATTCAGACAGATTCATTCATGATGTCAGAGATTTATACCTCTCCAGTGAGC 2180
DB 1178 AGAATTTATTCAGACAGATTCATTCATGATGTCAGAGATTTATACCTCTCCAGTGAGC 1237
QY 2181 TAGCAAGTATTCAGAGAGTGCAGCAAAACAAATAGTATACCTAGTACTTTAACT 2240
DB 1238 TGGCAAAATATTCAGAGAGTGCAGCAAAACAAATAGTATACCTAGTACTTTAACT 1297
QY 2241 GGAAGGAAGATTTCACTGTAAATCTTCCAGATTTTGGAGATCAGATGCTTGGCTT 2300
DB 1298 GGAAGGAAGATTTCACTGTAAATCTTCCAGATTTTGGAGATCAGATGCTTGGCTT 1357
QY 2301 GCGATCATGTGAAAAGGATCAAGATATATAGTCTGTGTAATTTAGAGAAATGCTTT 2360
DB 1358 GCGATCATGTGAAAAGGATCAAGATATATAGTCTGTGTAATTTAGAGAAATGCTTT 1417
QY 2361 GGAATTTAAATTTTATCATCTTGCACACTTGAATTAATTTGAGAGAGATATCAATCCAA 2420
DB 1418 GGAATTTAAATTTTATCATCTTGCACACTTGAATTAATTTGAGAGAGATATCAATCCAA 1473
QY 2421 GATTTGAGATTAAGAGAGATGCAACATATCTTTTGTGTGCAATTTATTTTATAC 2480
DB 1474 GATTTGAGATTAAGAGAGATGCAACATATCTTTTGTGTGCAATTTATTTTATAC 1528
QY 2481 CCTCTCTAGGATTAAGATGATATTTTGTGTGCAATTTTAAAGCTCAGCATAGCAATC 2540
DB 1529 CCTCTCTAGGATTAAGATGATATTTTGTGTGCAATTTTAAAGCTCAGCATAGCAATC 1588
QY 2541 ATTCATTTGCTTTTAAATATCTGATATACCTGATATATGATGATGCACTGGAGGATATTT 2600
DB 1589 ACTCAATTTGCTTTTAAATATCTGATATATGATGATGATGCACTGGAGGATATTT 1647

QY	2601	ATTCTTCATTATTCATTGTAACATTCCTTTTTCACATTTTGTAGTGTCCATTAATGTA	2660
DB	1648	ATTTCATATCTCATTTGTGACAGATTCGTGTTTACATTTATGAGTGTAAAGTA	1707
QY	2661	ACCTTGCTGTGATTTATTTTCCACAGTACAGCTGTTTAATCTATTGGGAAATGA	2720
DB	1708	ACTTTGTGATTT---TAATGTTTCCACATTCATCACTTTCACTTCTGGAAATGA	1764
QY	2721	AAATGCACATCTTAAAGTAGAA-AAATTTTCACTAATGATTAATCACTGATGTTCCAC	2779
DB	1765	AAATGCATATTTCTAAATTAAGAGAGCTTTTGCCAAAGTATGAATTTGATTTCTAA	1824
QY	2780	TTTGCACTACTATTAACAGAGAGAA	2804
DB	1825	CTTGTATATGCTTAACAAAAGGA	1849
RESULT 7			
AF345993		1128 bp	mRNA
LOCUS			
DEFINITION	Rattus norvegicus alpha1,3-fucosyltransferase IX (FucT) mRNA.		
ACCESSION	AF345993		
VERSION	AF345993.1	GI:13242183	
KEYWORDS			
SOURCE			
ORGANISM	Norway rat. Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 1128) Baboval, T., Hemion, T., Kinnally, E. and Smith, F.I. Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX) and comparison of the expression of Fuc-TIV and Fuc-TIX genes during rat postnatal cerebellum development J. Neurosci. Res. 62 (2), 206-215 (2000)		
JOURNAL	11020213		
REFERENCE	2 (bases 1 to 1128) Smith, F.I. and Baboval, T. Direct Submission Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver Center, 200 Trapele Rd., Waltham, MA, USA		
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source	1..1128 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" 1..1128 /gene="FucT9" 49..1128 /gene="Fuc9" /codon_start=1 /product="alpha1,3-fucosyltransferase IX" /protein_id="AAK16591.1" /db_xref="GI:13242184" /translation="MTSTKGLRPFLLVCIIICFVACLLIYIKPNSWVSPMESA SSYLKKNFESKRTIDFNETITLWVWFGGTEDLTSCOMENIOGCHTTRSLYRK SHAVLIHBDLSMDLNLDPQARPPQKIMNLESPTTHPOKSGIEHFNITLHYR DSIDQVYDGLITVTSPPFEVSEKREKYLQCMVNNMDEHARVYKNEHSEITLHYR GOAPGEVNDKNIPIITSNCKRYLESFENSJHKDVTIKIYNATLASVYVYGPSREN YENITPADSFTHVEDPNSPELAKTLKLVDAKNNKLITFTNMKDDTVNLPRFMESHA CLACDHRKROEYKSYGNLEKFM"		
BASE COUNT	324 a 256 c 228 g 320 t		
ORIGIN			
Query Match	33.2%;	Score 937.6;	DB 10; Length 1128;
Best Local Similarity	91.4%;	Pred. No. 3.9e-155;	
Matches 994; Conservative	0;	Mismatches 94;	Indels 0; Gaps 0;
QY	1281	GAATAATATGACATCAACATCCAAAGAAATCTTGGCCATTTTATTTGCTGCAATTA	1340
DB	41	GAATAATATGACCTCAACATCCAAAGCAATCTTGGCCATTTTATTTGCTGCAATTA	100
QY	1341	TCTGGGGCTGTTTCATGGCATGTCTTCTCATTTTACATCAAAACCTACAGAGCTGATCT	1400
DB	101	TCTGGGGCTGCTGTGGCATGTCTACTATATCTACATCAAGCCACCAAGCGTGGTCT	160
QY	1401	TCAGTCCAAAGGAATTCAGCCAGCGCTGTGTGAAAAATGAAAAATCTTTTCCACCAAAA	1460
DB	161	TCAGTCCAAAGGAGTCTGCAAGCTGTGTGAAAAATGAAAAATTTCTCTCCGAGAAA	220
QY	1461	CTGATTTATTTTAATGAACACTATCTTGTGTGTGGTGTGGCCATTTGGGCAACCTTG	1520
DB	221	CTGATTTATTTTAATGAACACTATCTTGTGTGTGGTGTGGCCATTTGGGCAACCTTG	280
QY	1521	ACCTTACATTCGTCGAAGCAATGTCACATCCAAAGATGATGCAATCCCAAGGACCGTT	1580
DB	281	ACCTTACATTCGTCGAAGCAATGTCACATCCAAAGGATGATGCAATCCCAAGGACCG	340
QY	1581	CAGTGTACACAAATCCCATGCAAGTCTGTATCATTCACGAGACATCATGTTGGAGTCTA	1640
DB	341	CGTGTACACAAATCCCATGCAAGTCTGTATCATTCACGAGACATCATGCTGGAGTCTA	400
QY	1641	CAATTTACCTGACAGAGTAGGCCACCTTCAGAAATGATTTGGATTTGGAAAT	1700
DB	401	CTAATTTACCTGACAGAGTAGGCCACCTTCAGAAATGATTTGGATTTGGAAAT	460
QY	1701	CACCAACTCACACTCCCAAAAGAGTGGCATTTGAGCACTTGTTAACCTGACTGACTT	1760
DB	461	CACCAACTCACACCCCAAAAGAGTGGCATTTGAGCACTTGTTAACCTGACTGACTT	520
QY	1761	ACGGCGGATTCAGATATTCACAGTGGCTTATGGCTTGTGACGTTAAGCAAAATCCCT	1820
DB	521	ATCGCGGATTCAGATATTCACAGTGGCTTATGGCTTGTGACGTTAAGCAAAATCCCT	580
QY	1821	TGCTGTGTAAGTGCACAGCAAGAAATGTGTGTCTGCTGGTGTGATTAATGCAATC	1880
DB	581	TGCTGTGTAAGTGCACAGCAAGAAATGTGTGTCTGCTGGTGTGATTAATGCAATC	640
QY	1881	CTGACATGCCAGATCAAGTATTTACATGAGCTAAGCAAAAGCAATTCATACCT	1940
DB	641	CCGAGATGCCAGGCTCAAGTATTTACATGAGCTAAGCAAAAGCAATTCATACCT	700
QY	1941	ACGGGCAAGCATTTGGAATATGTCATGATTTAAATTTGATTCCTCCATATCTGCTT	2000
DB	701	ATGGGCAAGCATTTGGAATATGTCATGATTTAAATTTGATTCCTCCATATCTGCTT	760
QY	2001	GTAATTTATCTTCTCTTGAATTAATCAATCCACAGAGATTAATCAATCCAGAAATG	2060
DB	761	GTAATTTATCTTCTCTTGAATTAATCAATCCACAGAGATTAATCAATCCAGAAATG	820
QY	2061	ACAATGCTTTTCTGCTGCTGTCTGTACCTGTGTTCTGGGACCATTTAGGAAATATG	2120
DB	821	ACAATGCAATTTTCTGCTGCTGTGTACCTGTGTTCTGGGACCATTTAGGAAATATG	880
QY	2121	AGAAATATTTCCAGACATATTCATTCATGTGGAATATTAATCTCCAGTAGC	2180
DB	881	AGAAATATTTCCAGACATATTCATTCATGTGGAATATTAATCTCCAGTAGC	940
QY	2181	TGCAAGATCTGAGAGAGTGCACAAAACAAATTAATTAATCTTACTTAACT	2240
DB	941	TGCAAGATCTGAGAGAGTGCACAAAACAAATTAATTAATCTTACTTAACT	1000
QY	2241	GGAGGAGAGATTTCACTGTAAATCTTCCACGATTTTGGGAATCAGATGATTTGGCTT	2300
DB	1001	GGAGGAGAGATTTCACTGTAAATCTTCCACGATTTTGGGAATCAGATGATTTGGCTT	1060
QY	2301	GGATCATGTGAAAAGGATCAAGAAATATAGTCTGTGTGTAATTTAGAAATGATTTT	2360
DB	1061	GGATCATGTGAAAAGGATCAAGAAATATAGTCTGTGTGTAATTTAGAAATGATTTT	1120
QY	2361	GGAAATTA 2368	
DB	1121	GGAAATTA 1128	

RESULT 8
AB035906 1707 bp DNA linear VRI 05-JAN-2002
LOCUS AB035906
DEFINITION Gallus gallus gene for CFU9, complete cds.
ACCESSION AB035906
VERSION AB035906.1 GI:18146865
KEYWORDS CFU9.
SOURCE Gallus gallus brain DNA.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (sites) Kaneo,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
The evolutionary history of glycosyltransferase genes
Unpublished
2 (bases 1 to 1707)
Kaneo,M., Saitou,N. and Kitano,T.
Direct Submission
Submitted (17-DEC-1999) Mlxa Kaneo, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Yata 111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6789)
COMMENT
Sequence updated (29-Feb-2000).
FEATURES
source
1..1707
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="brain"
268..1347
/gene="CFUc-TIX"
268..1347
/gene="CFUc-TIX"
/codon_start=1
/product="CFU9"
/protein_id="BAB82489.1"
/db_xref="GI:18146865"
translation="MSTSKSGIFRPFIVFVIGCFMFLIYIKPSSWISNPVESA
SSVLKMSKSFSSKDNFNETILVWLMPGQFEDLSCOTFNINCHLIDRSYLNK
SHAVLIHHDIMDLNLPQARPEPKIMNLSPTPKSGIEHLEFNTLYRR
DSDIQVPGSMVGTGATFYEVSKENLVCMVSNMNPBARVKKYINELSKYIEIHTY
GOAGDYVNDKLIPIISCKRYLSPENSINHDIYTEKLYNLLGASPVYVGPSEN
YENYIPDSFIHEDFLSPRELAETILMDKNKMYLSTFNKRF SVHLKFWESHA
CLACHVNRHQRKISGNLEKWFVN"
BASE COUNT 517 a 338 c 329 g 521 t 2 others
ORIGIN
Query Match 29.2%; Score 823.2; DB 5; Length 1707;
Best Local Similarity 76.4%; Pred. No. 4.4e-135;
Matches 1068; Conservative 1; Mismatches 309; Indels 20; Gaps 4;
QY 1268 CTTCTATTTCGTAGAAAATTTATGACATCAACATCCAAAGAAATTCCTGCGCATTTTAA 1327
DB 247 CACTGCGCTGGCAAAATTTATGACATCAACATCTAAAGAAATTTTCCGCCATTTT 306
QY 1328 ATTGTCGATTAATCTCGGCGTGTTCATGCGCATGTCTTCTCATTTACATCAAACTACC 1387
DB 307 ATTGTCGATTAATCTCGGCGTGTTCATGCGCATTTACTAATTAATTAACCAACA 366
QY 1388 AACAGTGAATCTTCAGTCCAAATGGAATGACGACGCTCTGTGTGAATAAAGAACTTC 1447
DB 367 AGCAGTGAATCTTCATCTCGTGGAAATGACGACGCTCAAGTTTGAATAAAGAACTTC 426
QY 1448 TTTTCCACCAAAATGATTAATTTAATGAACATCAATTCCTGCTGGTGCGCATTT 1507
DB 427 TTTTCTCCAAAATGATTAATTTAATGAACATTAATTTTGGTTGGCTTGGCCATTT 486
QY 1508 GGGCAGACCTTTGACCTTACATCTGCGCAAGCAATGTTCAACATCCAAAGATGCACTTC 1567
DB 487 GGCACAGATTCGATTAATCAATCTGCGCAAGCAATGTTCAACATCCAAAGATGCACTTC 546
QY 1568 ACACAGGACCTTCATCTGACCAAAATCCCATGAGTTCTGATCATCAACGAGCATC 1627
DB 1626 TTTTACATTTTCTGTGG 1643

DB 547 ACTATGACCGCTCACTAATTAACAAGTCCATGACGTTCTCATCATCAGGAGCAT 606
QY 1628 AGTTGGATCTGCAAAATTTTACTCAGCAAGTACGACCAACCTTCCAGAAATGATTTGG 1687
DB 607 AACTGGATCTGCAATTAATTTACTCAGCAAGTACGACCAACCTTCCAGAAATGATTTGG 666
QY 1688 ATGAATTTGGAATCACAACCTCAGCTCCCAAAAGAGTGGCATGAGCATTTGTTAAC 1747
DB 667 ATGAATTTGGAATCACAACCTCAGCTCCCAAAAGAGTGGCATGAGCATTTGTTAAC 726
QY 1748 CTGACCTGACCTAAGCGCGGTGATTAAGATATCCAAAGTGGCTTATGCTTTCAGCGTA 1807
DB 727 CTGACCTGACCTAAGCGCGGTGATTAAGATATCCAAAGTGGCTTATGCTTTCAGCGTA 786
QY 1808 AGCACAATTCCTCTGTTTGAAGTGCACAGCAAGAAATTTGGTGGCTTGGT 1867
DB 787 GGTACAGGCGCTTTCATATATGAGTCCAGTAAGTAAAGAAATTTGGTGGCTTGGT 846
QY 1868 AGTAAGTGAACCTTGAGCATGCGCAGATGCTCAAGTATTAAGTAAAGCAATGAT 1927
DB 847 AGTAAGTGAACCTTGAGCATGCGCAGATGCTCAAGTATTAAGTAAAGCAATGAT 906
QY 1928 GAAATGCAATCAGGAGGCAAGCATTTGAGAAATATGCAATGATTAATTTGATTCCT 1987
DB 907 GAAATGCAATCAGGAGGCAAGCATTTGAGAAATATGCAATGATTAATTTGATTCCT 966
QY 1988 ACCATATCTGCTTGAATTTATCTTCTCTTGAATTTCAATCCACAAGATTAATCAT 2047
DB 967 ACATATCTGCAATTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTT 1026
QY 2048 ACGAAAAGCTTATCAATGCTTTGCGTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 2107
DB 1027 ACGAAAAGCTTATCAATGCTTTGCGTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1086
QY 2108 AGGAAAACATATGAGATTAATATATCCAGAGATTTCAATTAATTTGAGATTTAAC 2167
DB 1087 AGGAAAACATATGAGATTAATATATCCAGAGATTTCAATTAATTTGAGATTTAAC 1146
QY 2168 TCTCCAGTACGCTACCAAGTATGGAAGAAAGTGCACAAAACAAATGATTAATCAT 2227
DB 1147 TCCCAAGAGAGCTTCCAGAAATATCTTCTGATGCTTGCACAAAACAAATGATTAAT 1206
QY 2228 AGTTACTTAACTGAGGAGGATTTCACTGTAATCTTCAACATTTTGGGAATCAT 2287
DB 1207 AGTTACTTAACTGAGGAGGATTTCACTGTAATCTTCAACATTTTGGGAATCAT 1266
QY 2288 GCATGTTGCGCTGCGATATGTAAGAAAGCATCAAGATTAATGCTTGTGATTTTA 2347
DB 1267 GCATGTTGCGCTGCGATATGTAAGAAAGCATCAAGATTAATGCTTGTGATTTTA 1326
QY 2348 GAGAAATGCTTTGGAATTA--AATTTTCAATCACTGACACCTGATTAATTTTGA 2405
DB 1327 GAGAAATGCTTTGGAATTAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1386
QY 2406 TGAGATATCAATCAAGTATTAAGATTAAGAA-----GAGATGCAACATA 2449
DB 1387 GATCTTGTAGTCAATGAGGATTAAGAAATTAAGAAATTAAGGATTAAGGATTAAG 1446
QY 2450 CTACTTTTGTGACAAATTTATTTTATCACCTCTCTGAGGATGATTAATTTGGT 2509
DB 1447 ACAGAACTTACGTCAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 1506
QY 2510 GGAGATTTTAAAGCTGCAATGAGCAATCAAT--TCCATTCGTTTAAATTAATCTGTA 2568
DB 1507 TAAGAGATTTTAAAGCTGCAATGAGCAATCAAT--TCCATTCGTTTAAATTAATCTGTA 1566
QY 2569 TATACCTAATTAATGCACTGAGAGTAAATTAATTTCTTCAATTAATTTGTAACATGTC 2628
DB 1567 CATTAATTAATGCACTGAGAGTAAATTAATTTCTTCAATTAATTTGTAACATGTC 1625
QY 2629 TTTTTCACATTTTGTAG 2646
DB 1626 TTTTTCACATTTTGTAG 1643

RESULT 9
AB035905 1080 bp DNA linear VRT 05-JAN-2002
DEFINITION Xenopus laevis gene for xFUT9, complete cds.
ACCESSION AB035905
VERSION AB035905.1 GI:18146863
KEYWORDS xFUT9.
SOURCE Xenopus laevis DNA.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
Xenopodidae; Xenopus.
REFERENCE
1 (sites)
AUTHORS Kaneo,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL Unpublished
AUTHORS Kaneo,M., Saitou,N. and Kitano,T.
TITLE 2 (bases 1 to 1080)
JOURNAL Direct Submission
Submitted (17-DEC-1999) Mka Kaneo, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Yata 111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6789)
Sequence updated (29-Feb-2000).
COMMENT Location/Qualifiers
FEATURES
source
1..1080
/organism="Xenopus laevis"
/db_xref="taxon:8355"
1..1080
/gene="xFuc-TIX"
1..1080
/gene="xFuc-TIX"
/codon_start=1
/product="xFUT9"
/protein_id="BAB82488.1"
/db_xref="GI:18146864"
/translation="MTSASKGIILRAFLAVSGVILICFNICILLYVPTNNMISPIESA
HSVILKHNFSKONYNEVLIWVPGFOTELSCOTLNIYCHLITDRALYNK
SHAVLHHRDISMDLINLPOLRPPPKMIMLESPPHPPKSGIEHLFDITLTYR
DSIDIOYGFEMVYSTRKPRDEPVSKDKYCVSWNNPKNLILAGSPYILYIGPREN
GOAFGEYSDEKSLPTASCEFYSLSPENSHKRYTEKXNLLAGSPYILYIGPREN
YENYIDASFTYDEPLSPRELADHMLMKDKDEQLISTYNNKAFYVAMPHMESHA
CLACDHYKRDHYKSVGNLEKERN"
BASE COUNT 332 a 231 c 202 g 315 t
ORIGIN
Query Match 24.0%; Score 676.8; DB 5; Length 1080;
Best Local Similarity 76.7%; Pred. No. 2.5e-109;
Matches 828; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

Db 301 AACAACTCATGCTGCTCTTCTTACACAGACATATGATGATCTATTAACCTG 360
1649 CCTCAGCAAGCTAGCCACCCCTTCCAGAAATGATTTGATTAATCCACCACT 1708
Db 361 CCCCAAGAGCTGAGGCCACCTTCCAGAAATGATTTGATTAATCCACCACT 420
1709 CACACTCCCAAGAGCTGAGCCCTTCCAGAAATGATTTGATTAATCCACCACT 1768
Db 421 CATTACCCCAAGAGCTGAGCCCTTCCAGAAATGATTTGATTAATCCACCACT 480
1769 GATTACATATTCAGAGCTGAGCCCTTCCAGAAATGATTTGATTAATCCACCACT 1828
Db 481 GACTAGATATTCAGAGCTGAGCCCTTCCAGAAATGATTTGATTAATCCACCACT 540
1829 GAAGTCCCAAGAGCTGAGCCCTTCCAGAAATGATTTGATTAATCCACCACT 1888
Db 541 GAGGTCCCAAGAGCTGAGCCCTTCCAGAAATGATTTGATTAATCCACCACT 600
1889 GCCAGAGCTAGATTTCAATGAGCTAGAGCAAGATTAATCCACCACTAGGCA 1948
Db 601 GCCAGAGCTAGATTTCAATGAGCTAGAGCAAGATTAATCCACCACTAGGCA 660
1949 GCAATTTGAGAAATGATTAATGATTAATGATTTGATTAATCCACCACTAGGCA 2008
Db 661 GCTTTGAGAAATGATTAATGATTAATGATTTGATTAATCCACCACTAGGCA 720
2009 TATCTTCTTCTTGAATTAATGATTAATGATTTGATTAATCCACCACTAGGCA 2068
Db 721 TATCTTCTTCTTGAATTAATGATTAATGATTTGATTAATCCACCACTAGGCA 780
2069 TTTCTGAGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 2128
Db 781 TTTCTGAGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 840
2129 ATTCCAGAGATTAATGATTAATGATTTGATTAATGATTTGATTAATCCACCACT 2188
Db 841 ATTCCAGAGATTAATGATTAATGATTTGATTAATGATTTGATTAATCCACCACT 900
2189 TATCTGAGAGCTGATTAATGATTTGATTAATGATTTGATTAATCCACCACT 2248
Db 901 CATCTTTTATGCTGATTAATGATTTGATTAATGATTTGATTAATCCACCACT 960
2249 GATTTCTGATTAATGATTTGATTAATGATTTGATTAATCCACCACTAGGCA 2308
Db 961 CATTTTACAGCTATTAATGATTTGATTAATGATTTGATTAATCCACCACT 1020
2309 GTGAAAGCATCAAGATTAATGATTTGATTAATGATTTGATTAATCCACCACT 2368
Db 1021 GTGAAAGCATCAAGATTAATGATTTGATTAATGATTTGATTAATCCACCACT 1080

RESULT 10
AC100674 71396 bp DNA linear HTG 22-NOV-2001
LOCUS AC100674
DEFINITION Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC100674
VERSION AC100674.1 GI:17048040
KEYWORDS HTG; HTGS; PHASED.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 71396)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Unpublished
JOURNAL 2 (bases 1 to 71396)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguski,M., Bouckgeit,B.,
Brown,A., Camarata,J., Campione,A., Chang,J., Chazaro,B.,
Choepe,J., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,

TITLE
JOURNAL
COMMENT

Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,
Grinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Labrecque, K.,
Lamaze, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McKenney, P., McPherson, K., McPherson, K.,
Meneses, L., Milova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliyer, J., Peterson, K., Phunkhang, P., Pierle, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Strauss, K., Subramanian, A., Talamas, J., Testaye, S., Theodores, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zaloun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smilt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L15932

Center clone name: 167_K_24

* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 789 888: contig of 788 bp in length
* 889 1702: contig of 814 bp in length
* 1703 1802: contig of 100 bp
* 1803 2642: contig of 840 bp in length
* 2643 2742: contig of 100 bp
* 2743 3637: contig of 895 bp in length
* 3638 3737: contig of 100 bp
* 3738 4635: contig of 898 bp in length
* 4636 4735: contig of 100 bp
* 4736 5657: contig of 922 bp in length
* 5658 5757: contig of 100 bp
* 5758 6692: contig of 935 bp in length
* 6693 6792: contig of 100 bp
* 6793 7703: contig of 911 bp in length
* 7704 7803: contig of 100 bp
* 7804 8742: contig of 939 bp in length
* 8743 8842: contig of 100 bp
* 8843 9787: contig of 945 bp in length
* 9788 9887: contig of 100 bp
* 9888 10819: contig of 932 bp in length
* 10820 10919: contig of 100 bp
* 10920 11808: contig of 889 bp in length
* 11809 11908: contig of 100 bp
* 11909 12830: contig of 922 bp in length
* 12831 12930: contig of 100 bp
* 12931 13847: contig of 917 bp in length
* 13848 13947: contig of 100 bp
* 13948 14886: contig of 939 bp in length
* 14887 14986: contig of 100 bp

14987 15911: contig of 925 bp in length
* 15912 16011: gap of 100 bp
* 16012 16903: contig of 892 bp in length
* 16904 17003: gap of 100 bp
* 17004 17906: contig of 903 bp in length
* 17907 18006: gap of 100 bp
* 18007 18931: contig of 925 bp in length
* 18932 19031: gap of 100 bp
* 19032 19945: contig of 914 bp in length
* 19946 20045: gap of 100 bp
* 20046 20951: contig of 906 bp in length
* 20952 21051: gap of 100 bp
* 21052 21949: contig of 898 bp in length
* 21950 22049: gap of 100 bp
* 22050 22950: contig of 901 bp in length
* 22951 23050: gap of 100 bp
* 23051 23945: contig of 895 bp in length
* 23946 24045: gap of 100 bp
* 24046 24955: contig of 910 bp in length
* 24956 25055: gap of 100 bp
* 25056 25980: contig of 925 bp in length
* 25981 26080: gap of 100 bp
* 26081 27052: contig of 972 bp in length
* 27053 27152: gap of 100 bp
* 27153 28067: contig of 915 bp in length
* 28068 28167: gap of 100 bp
* 28168 29076: contig of 909 bp in length
* 29077 29176: gap of 100 bp
* 29177 30093: contig of 917 bp in length
* 30094 30193: gap of 100 bp
* 30194 31099: contig of 906 bp in length
* 31100 31199: gap of 100 bp
* 31200 32097: contig of 898 bp in length
* 32098 32197: gap of 100 bp
* 32198 33113: contig of 916 bp in length
* 33114 33213: gap of 100 bp
* 33214 34125: contig of 912 bp in length
* 34126 34225: gap of 100 bp
* 34226 35133: contig of 908 bp in length
* 35134 35233: gap of 100 bp
* 35234 36150: contig of 917 bp in length
* 36151 36250: gap of 100 bp
* 36251 37171: contig of 921 bp in length
* 37172 37271: gap of 100 bp
* 37272 38159: contig of 888 bp in length
* 38160 38259: gap of 100 bp
* 38260 39156: contig of 897 bp in length
* 39157 39256: gap of 100 bp
* 39257 40188: contig of 932 bp in length
* 40189 40288: gap of 100 bp
* 40289 41210: contig of 922 bp in length
* 41211 41310: gap of 100 bp
* 41311 42222: contig of 912 bp in length
* 42223 42322: gap of 100 bp
* 42323 43218: contig of 896 bp in length
* 43219 43318: gap of 100 bp
* 43319 44233: contig of 915 bp in length
* 44234 44333: gap of 100 bp
* 44334 45246: contig of 913 bp in length
* 45247 45346: gap of 100 bp
* 45347 46261: contig of 915 bp in length
* 46262 46361: gap of 100 bp
* 46362 47279: contig of 918 bp in length
* 47280 47379: gap of 100 bp
* 47380 48276: contig of 897 bp in length
* 48277 48376: gap of 100 bp
* 48378 49282: contig of 906 bp in length
* 49283 49382: gap of 100 bp
* 49383 50305: contig of 923 bp in length
* 50306 50405: gap of 100 bp
* 50406 51314: contig of 909 bp in length
* 51315 51414: gap of 100 bp
* 51415 52335: contig of 921 bp in length

```

* 52336 52435: gap of 100 bp
* 52436 52535: contig of 728 bp in length
* 53164 53263: gap of 100 bp
* 53264 54210: contig of 947 bp in length
* 54211 54310: gap of 100 bp
* 54311 55199: contig of 889 bp in length
* 55200 55299: gap of 100 bp
* 55300 56212: contig of 913 bp in length
* 56213 56312: gap of 100 bp
* 56313 57287: contig of 975 bp in length
* 57288 57387: gap of 100 bp
* 57388 58245: contig of 858 bp in length
* 58246 58345: gap of 100 bp
* 58346 59236: contig of 891 bp in length
* 59237 59336: gap of 100 bp
* 59337 60272: contig of 936 bp in length
* 60273 60372: gap of 100 bp
* 60373 61289: contig of 917 bp in length
* 61290 61389: gap of 100 bp
* 61390 62299: contig of 910 bp in length
* 62300 62399: gap of 100 bp
* 62400 63322: contig of 923 bp in length
* 63323 63422: gap of 100 bp
* 63423 64336: contig of 914 bp in length
* 64337 64436: gap of 100 bp
* 64437 65354: contig of 918 bp in length
* 65355 65454: gap of 100 bp
* 65455 66352: contig of 898 bp in length
* 66353 66452: gap of 100 bp
* 66453 67362: contig of 910 bp in length
* 67363 67462: gap of 100 bp
* 67463 68370: contig of 908 bp in length
* 68371 68470: gap of 100 bp
* 68471 69379: contig of 909 bp in length
* 69380 69479: gap of 100 bp

```

Query Match 22.8%; Score 644; DB 2; Length 71396;
 Best Local Similarity 75.1%; Pred. No. 6,7e-104;
 Matches 784; Conservative 0; Mismatches 247; Indels 13; Gaps 4;

```

OY 1215 AATATATATATGATTTTCCTTCATTCCTCAACCGCTACCTCCCTTCCTTCCTTCATTT 1274
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65311 AAAAGAGTAGTGATGATGAATTAATTAATGAATATACCAANNNNNNNNNNNNNNNN 65370
OY 1275 TCGTAGGAAAAATTTAGCATCAACATCCAAAGAAATTCGCGCCATTTTAAATGCT 1334
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65371 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 65430
OY 1335 GCATTAATCTGGGCTGTTTCATGCGATGCTTCTCATTTACATCAACCTACCAAGCT 1394
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65431 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 65490
OY 1395 GGATCTTCAGTCCAAATGGAATCAGCGAGCTGTCGGAAGAAATGAAAACTTCCTTCCA 1454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65491 CGGTCTTCAGTCCAAATGGAATGAGTCTGCAAGTCTGCGGAATGAAAAATTTCTCTCCA 65550
OY 1455 CCAAACTGATTTATTTATGAACATCTATCTGAGTGGGGGGGGGCGCCATTGGGCGGA 1514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65551 CAAAACTGATTTATTTAAAGAAATCAGATCTGCTTGGTATGGCATTTGGGCGGA 65610
OY 1515 CCTTGAAGCTTACATCTGCGCAAGCAATGTTCAACATCCAAAGATGCGCTCACACGG 1574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65611 CCTTGAAGCTTACATCTGCGCAAGCAATGTTCAATATCCAAAGGATGCGCTCACACAG 65670
OY 1575 ACCGTTCACTGTACAACAATCCATCAGTTCATGATCCATCCAGAGACATCACTTGGG 1634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65671 ACCGTTCACTGTACAACAATCCATCAGTTCATGATCCATCCAGAGACATCACTTGGG 65730
OY 1635 ATCTGACAAATTTACCTAGCAGAGCTAGCGACCCCTTCGAAATGAGATTTGATGAAATT 1694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65731 ATCTGACAAATTTACCTAGCAGAGCTAGCGACCCCTTCGAAATGAGATTTGATGAAATT 65790
OY 1695 TGAATACCACTACACACTCCCAAAAAGATGAGCATTTGTTTAACTGACTC 1754

```

```

Db 65791 TAAAGTACCCACACTACACCCCAAAAAGATGGCATTAACATGTTCAACTGACTC 65850
OY 1755 TGAATACCCACACTACACCCCAAAAAGATGGCATTAACATGTTCAACTGACTC 1814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65851 TAACTATTCGCGGTATTCATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 65910
OY 1815 ATCCCTGCTGTTTAAATGCGCAAGCAAGAAATGGTGTGCTGGTGTGAGTAATC 1874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65911 ATCCCTGCTGTTTAAATGCGCAAGCAAGAAATGGTGTGCTGGTGTGAGTAATC 65970
OY 1875 GGAACCTGAGATCCGAGATCAAGATTAATGATTAAGTAAAGATTAAGTAAATC 1934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65971 GGAACCTGAGATCCGAGATCAAGATTAATGATTAAGTAAAGATTAAGTAAATC 66030
OY 1935 ATACCTAGCGGCAAGCATTTGAGATATGTCATATTAATAAATTTGATTCCTACATAT 1994
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66031 ACACCTATGCGCAAGCATTTGAGATATGTCATATTAATAAATTTGATTCCTACATAT 66090
OY 1995 CTGCTGTAAATTTATCTTCTCTTGAATAATTCATTCACAAAGATTCACAGGAAA 2054
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66091 CTACTGTAAATTTATCTTCTCTTGAATAATTCATTCACAAAGATTCACAGGAAA 66146
OY 2055 AGCTATACAAATCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66147 TCCACATGCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 66201
OY 2115 ACTATGAGATTAATTTATTCAGACAGATTCATTCATTCATTCATTCATTCATTCATTCAT 2171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66202 ACTATGAGATTAATTTATTCAGACAGATTCATTCATTCATTCATTCATTCATTCATTCAT 66261
OY 2172 CCAATGAGTACCAAAATGATCTGAAGAGTGCACAAAACATTAATTT-ATACCTAGT 2230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66262 CCAATGAGTACCAAAATGATCTGAAGAGTGCACAAAACATTAATTT-ATACCTAGT 66321
OY 2231 TACTTAACTGGAGAGAGATTTTC 2254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66322 TACTTAACTGGAGAGAGATTTTC 66345

```

RESULT 11
 G55571/c 561 bp DNA linear STS 30-MAR-2000
 LOCUS SHGC-100926 Human Homo sapiens STS genomic, sequence tagged site.
 DEFINITION SHGC-100926 Human Homo sapiens STS genomic, sequence tagged site.
 ACCESSION G55571
 VERSION G55571.1 GI:6120890
 KEYWORDS STS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 561)
 AUTHORS Olivier, M. and Cox, D. R.
 TITLE Unpublished, Olivier, M., Cox, D. R. (2000)
 JOURNAL Unpublished
 COMMENT

Contact: Michael Olivier, David R. Cox
 Stanford Human Genome Center
 Stanford University School of Medicine
 4005 Miranda Ave., 2nd Fl., Palo Alto, CA 94025, USA
 Tel: (650) 320-5800
 Fax: (650) 320-5801
 Email: Olivier@shgc.stanford.edu
 Primer A: CAGAACTTCTCTCAAGGCCAGCA
 Primer B: TAACGAATTCATCGAACACCA
 STS size: 302
 PCR Profile:
 Initial incubation: 95 degrees C for 10 minutes
 Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9700

Protocol:

Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer:

MgCl₂: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

BAC ends sequenced at TIGR from the RPC11 BAC library. Designed and developed at the Stanford Human Genome Center.

FEATURES
 Location/Qualifiers
 1..561

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="6"

STS
 primer_bind 153..454
 primer_bind 153..175
 complement(432..454)

BASE COUNT 200 a 92 c 89 g 180 t
 ORIGIN

Query Match 19.8%; Score 557.8; DB 11; Length 561.

Best Local Similarity 99.6%; Pred. No. 2.2e-88; Mismatches 2; Indels 0; Gaps 0;

Matches 559; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 717 TGGCATTCTTTTACGATGATTAATGATTCGAACTGGAAGTAAATTTAT 776
 DB 561 TGGCATTCTTTTACGATGATTAATGATTCGAACTGGAAGTAAATTTAT 562
 QY 777 AAGCAGAAAAACAAACATACGTTTAAAACAGTTTACTTTCTAAAGTAATTCAT 836
 DB 501 AAGCAGAAAAACAAACATACGTTTAAAACAGTTTACTTTCTAAAGTAATTCAT 442
 QY 837 CTGACACCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 896
 DB 441 CTGACACCAAAAGTATGATGATGATGATGATGATGATGATGATGATGAT 382
 QY 897 GGGAGATGTTTACAAATCTGATGTTTACAAATCTGATGTTTACAAATCTGAT 956
 DB 381 GGGAGATGTTTACAAATCTGATGTTTACAAATCTGATGTTTACAAATCTGAT 322
 QY 957 GCGATGACAGTCTAGGTTGGCCACTTTTCTACAGCAATTAAGTAGACGTTAAG 1016
 DB 321 GCGATGACAGTCTAGGTTGGCCACTTTTCTACAGCAATTAAGTAGACGTTAAG 262
 QY 1017 ATAGCATTTGAAAAAGCTTACTTGGTTGAAATATGCTGATCTTTAGAGAGTTTT 1076
 DB 261 ATAGCATTTGAAAAAGCTTACTTGGTTGAAATATGCTGATCTTTAGAGAGTTTT 202
 QY 1077 TGGGAAACTCTCTTTGAAAGCTGTTTCTGCTTGGAGAAAGTCTGTTCTCAAT 1136
 DB 201 TGGGAAACTCTCTTTGAAAGCTGTTTCTGCTTGGAGAAAGTCTGTTCTCAAT 142
 QY 1137 GCTTATATACACTGAAAGGAAATGCTGTTTACTTACTTCTTATCTCAATTTAT 1196
 DB 141 GCTTATATACACTGAAAGGAAATGCTGTTTACTTACTTCTTATCTCAATTTAT 82
 QY 1197 GATTTAATATATATCAATATATATATATATATATATATATATATATATATAT 1256
 DB 81 GATTTAATATATATCAATATATATATATATATATATATATATATATATATAT 22
 QY 1257 CTCTGCTCTTCTCTAATTCG 1277
 DB 21 CTCTGCTCTTCTCTAATTCG 1

RESULT 12

AC100674/c

LOCUS AC100674 71396 bp DNA linear HTG 22-NOV-2001
 DEFINITION Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.
 AC100674
 VERSION 1
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Brown, A., Camarata, J., Campilongo, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, D., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand, P., Hago, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Meneus, L., Minova, T., Mienze, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhong, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Topham, K., Travers, M., Travalis, N., Triggillo, J., Vassiliev, J., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, L., and Zody, M.

TITLE

JOURNAL Direct Submission
 Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/xw/RepeatMasker.html

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L15932
 Center clone name: 167_K_24

* NOTE: This record contains 71 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 788: contig of 788 bp in length
 * 789 888: gap of 100 bp
 * 889 1702: contig of 814 bp in length
 * 1703 1802: gap of 100 bp
 * 1803 2642: contig of 840 bp in length
 * 2643 2742: gap of 100 bp
 * 2743 3637: contig of 895 bp in length
 * 3638 3737: gap of 100 bp
 * 3738 4635: contig of 898 bp in length
 * 4636 4735: gap of 100 bp

Db	30030	TTGGGATTTACATGCATGCGCTGGCATGCGACCATGTAAAAAGGCATCAAGATATATAGTC	29971
Qy	2335	TCCTGGTAATTTAGAGAAAATGGCTTTTGGAAATTAATAATTTTTCATCCACTTGCACACTTGGAT	2394
Db	29970	TCTGGTATATTTAGAGAAATGGTTTGGAAAT- AAAGTCTCATCTATCTGTACACA---GTG	29915
Qy	2395	AAATATTTTGATGAGATATATCATCCCAAGTGTAGGATTAAGAAAGAGATGCAACATCTACT	2454
Db	29914	AAAAAGATTAAAGAACATTAATTCAGTTTCAGGATTAAGAA-----AACACGGCTGCA	29860
Qy	2455	TTTGATGCACATTTATTTTATTCACACCCCTCTAGAGTAACGGTATATTTGGTGGAGA	2514
Db	29859	TTTGGGGCACCGCTTATATTTTCTGCGCCCTCGCTTGAGGACATGTATATTTGGTGA- A	29801
Qy	2515	TTTTTAAAAAGCTCAGCATGAGCAATCATTCATCGGTTTTTAAATTAATCCTGTATATACC	2574
Db	29800	TTTTTAAGAGATCAGAAATTAAGCAATCAGTCATCAATTTGGTTTAAATATATCCTGTAAATANG	29741
Qy	2575	TAAATATGTCAGCTGGAGAGTAATTTATCTTCATATATCATTTGTGTAACATTGCTTTTC	2634
Db	29740	TGATTAACGAGCACTGGA- AATATATTTATTTTGGCACACCTATTAATAAACATTTGTTTATC	29682
Qy	2635	ACATTTTGTAGTATGTCATATGTAATGAAGCTTGCGTTGATTAATGTTCCACACAGTAC	2694
Db	29681	A---TTTATAGTTACACGTAAAGTAATTTATGATT---TACTGTCTTCTACATCAATC	29628
Qy	2695	AGCTGTTTAAATCTAATTTGGGAAATGAAGATGACATCTTAAAGTANG- AAAATTTTCCAC	2753
Db	29627	AGATCTTTAACTATTTTGGGAGATAAAAATACATATCTCAAAAATATGAGGACCTTTTGC	29568
Qy	2754	TAAATATTAACAATGCTAGTCCACTTCCACTTGCATACTATATAACAGAGAGAA	2804
Db	29567	CAGATTTATATAAGTATAGATTCTAGCTTGATATATGCTAACAAAGAGAA	29517

RESULT	13				
LOCUS	AB023627				
DEFINITION	AB023627	1567 bp	DNA	linear	VMT 02-APR-1999
	Danio rerio zft1 gene for alpha(1,3)fucosyltransferase, complete cds.				
ACCESSION	AB023627				
VERSION	AB023627.1	GI:4587295			
KEYWORDS	alpha(1,3)fucosyltransferase.				
SOURCE	Danio rerio DNA.				
ORGANISM	Danio rerio				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.				
REFERENCE	1 (sites)				
AUTHORS	Kageyama,N., Natsuka,S. and Hase,S.				
TITLE	Molecular cloning and characterization of two zebrafish alpha(1,3)fucosyltransferase genes developmentally regulated in embryogenesis				
JOURNAL	J. Biochem. (1999) In press				
REFERENCE	2 (bases 1 to 1567)				
AUTHORS	Kageyama,N., Natsuka,S. and Hase,S.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate School of Science; MachikaneYama 1-1, Toyonaka, Osaka 560-0034, Japan (E-mail:natsuka@chem.sci.osaka-u.ac.jp, Tel:81-6-6850-5381, Fax:81-6-6850-5383)				
FEATURES					
SOURCE	Location/Qualifiers				
	1..1567				

```

gene      177.  .1352
          /gene="zFrl"
CDS       177.  .1352
          /gene="zFrl"
          /codon_start=1
          /product="alpha(1,3)fucosyltransferase"
          /protein_id="BAA6706.1"

```

Query Match	Best Local Similarity	Matches	Conservative	Score	DB	Length
1471	9.9%	58.3%	0	279.6	5	1567
461	58.3%	0	369	1.3e-39	12	2
1531	58.3%	0	369	1.3e-39	12	2
521	58.3%	0	369	1.3e-39	12	2
1591	58.3%	0	369	1.3e-39	12	2
581	58.3%	0	369	1.3e-39	12	2
1651	58.3%	0	369	1.3e-39	12	2
632	58.3%	0	369	1.3e-39	12	2
1711	58.3%	0	369	1.3e-39	12	2
692	58.3%	0	369	1.3e-39	12	2
1771	58.3%	0	369	1.3e-39	12	2
752	58.3%	0	369	1.3e-39	12	2
1828	58.3%	0	369	1.3e-39	12	2
812	58.3%	0	369	1.3e-39	12	2
1888	58.3%	0	369	1.3e-39	12	2
872	58.3%	0	369	1.3e-39	12	2
1948	58.3%	0	369	1.3e-39	12	2
932	58.3%	0	369	1.3e-39	12	2
2008	58.3%	0	369	1.3e-39	12	2
992	58.3%	0	369	1.3e-39	12	2
2068	58.3%	0	369	1.3e-39	12	2
1052	58.3%	0	369	1.3e-39	12	2
2128	58.3%	0	369	1.3e-39	12	2
1112	58.3%	0	369	1.3e-39	12	2
2188	58.3%	0	369	1.3e-39	12	2
1172	58.3%	0	369	1.3e-39	12	2
2248	58.3%	0	369	1.3e-39	12	2
1232	58.3%	0	369	1.3e-39	12	2
2308	58.3%	0	369	1.3e-39	12	2

DB	1292	TCTAAAGAAATACACATATATAGCTTTAAAGATCTGAAAGCGTGTTCGGGGTTA	1351
OY	2368	AAATTTTCATCAG	2381
DB	1352	ATCGCATTTCTTAC	1365
RESULT 14			
LOCUS	AB023628		
DEFINITION	AB023628 1832 bp DNA linear VRT 02-APR-1999		
ACCESSION	Dano rerio zFT2 gene for alpha(1,3)fucosyltransferase, complete cds		
VERSION	AB023628.1 GI:4587297		
KEYWORDS	alpha(1,3)fucosyltransferase.		
SOURCE	Dano rerio DNA.		
ORGANISM	Dano rerio		
REFERENCE	Ekarayota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Dano.		
AUTHORS	Kageyama, N., Natsuka, S. and Hase, S.		
TITLE	Molecular cloning and characterization of two zebrafish alpha(1,3)fucosyltransferase genes developmentally regulated in embryogenesis		
JOURNAL	J. Biochem. (1999) In press		
REFERENCE	2 (bases 1 to 1832)		
AUTHORS	Kageyama, N., Natsuka, S. and Hase, S.		
TITLE	Direct Submision		
JOURNAL	Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034, Japan (E-mail:natsuka@chem.sci.osaka-u.ac.jp, Tel:81-6-6850-5361, Fax:81-6-6850-5383)		
FEATURES	Location/Qualifiers		
source	1..1832		
gene	/organism="Dano, rerio"		
	/db_xref="taxon:7955"		
	332..1510		
	/gene="zFT2"		
	332..1510		
	/gene="zFT2"		
	/codon_start=1		
	/product="alpha(1,3)fucosyltransferase"		
	/protein_id="BA06707.1"		
	/db_xref="GI:4587298"		
	/translation="MDKTVADKIKRLAPRSKSTQKLIIVTFMILSVCIFFVYVNPNTTFKTPMGHIANSSCSESTLHAKMONEFCIKRNASVALPAPNPAKDQDIIIVIMAFPRVPEIDKDGLEYNIHGQLYDYSLLQKAGVPHHRMSDFQPPPPEDQKIMWNMESPTHCNQNDFLNDNRNVSSYNSRSDIPVYGLVATEQSKFTIPKDKLVCWISNFQGNHHRKHFYEFERAKIINVSYGFHNRINGDDYGNVSSCFYLSLFSNLSHRYETKLEPLNPLATGVVYVGPVRYEEFIPRDAPIVHDDPPPKELABHLLSLDNEOYKQEPFNMRKILVSKSYSGLEHSCRIDYLRKRYIAVTLKGMFWG"		
BASE COUNT	585 a	394 c	317 g 536 t
ORIGIN			
Query Match	8.3%; Score 234.6; DB 5; Length 1832;		
Best Local Similarity	55.5%; Pred. No. 1e-31;		
Matches 500; Conservative	0; Mismatches 389; Indels 12; Gaps 2;		
OY	1471	TAATGAACACATATCTGTGTGGGTGGGCAATTTGGGAGACCTTTGACTTACATC	1530
DB	619	TCAAGACATCATCGTGTGATCTGGATGGGCTCTTTCAGACTACTCTTACCTGAAAGA	678
OY	1531	CTGGCAACGAACTTCAACATCCAGAGATGCATCTCCACAGGAGCGTCACTGACAA	1590
DB	679	TTTGCGTTTGGATATTAACATCCAGCGCTGTCAATTAACACAGAGAAGTGTGATCCA	738
OY	1591	CAATCCAGACGCTCTGATCCACACCGACACATCAGTGGTGGATCTGCACAAATTACC	1650
DB	739	GAAGCGCATGAGATTATGTTTCCACACAGAGATATGAGC-----ACAGATTTC	789
OY	1651	TCAAGCAAGTCAGGCCACCTTCAGAAATGAGATTGGATGTAATTCACCAACTCA	1710

Db	790	GCACCCCGCGACGACGATTTCCAAAATGGATATGGTGGAAATATGATGATGCTCCACCCA	849
Qy	1711	CACCTCCCAAAAAGATGGGATGTAGCACTGTTTAACTTAACTTACCTGATCTGATACCCGCCCA	1770
Db	850	TTGCATTCAAAATGACCTTCTTAATGATCGGTTTAAATGTAGTACCTGCAAGTTACACAGSGA	909
Qy	1771	TTCAAGATTCGAAGTGCCTTATAGCTTCTTACGCGTAAGCA---AATCCCTTGCTTT	1827
Db	910	TTCCGATATTCAGTGTCTTACGCTGACCTGCTGTGATGCCAGAGAGACAGAAAAAGTT	969
Qy	1828	TGAAGTCCCAAGCAAGAAATTTGTGTGCTGGGTGTGAGTAACTGTGAACCTTGAGCA	1887
Db	970	CACCATCCCAAAAAGGATTAACTAATGTTGTGATAGTAAGTAACTTTCACAGGCGACCA	1029
Qy	1888	TGCCAGAGTCAAGTATTAACAATGACCTAAGCAAAAAGCATTTGAATCCATCCATCAGCGCA	1947
Db	1030	CAAGCGACACATTTCTTCTAGAGATTGGCGAAACACATTAATGTAGACACCTATGGAGG	1089
Qy	1948	AGCATTTGGAGATATGTCAATGATAAAAATTTGATTCACCATATATCTGCTGTAAAT	2007
Db	1090	GCATTTCAACATTCGATTAATGATGACGATTAATGAAATGCTGATCCAGTTGTAAT	1149
Qy	2008	TTATCTTTCCCTTTGAAAATTCATTCACCAAGATTAATCATCGGAAACATATACAAATGC	2067
Db	1150	CTACAGTCTCTTGAATCATTCATTCACAGAGCATTTTTCACGAGACAGCTGTCATCC	1209
Qy	2068	TTTTTGGCGGCGTGTGATCTGTTCTGTGGAGCATATCAGGAAAGCAATATGAGAAATTA	2127
Db	1210	TTTAGACATCTGGATAGTTCCTGTGTATTTGGCCGCCGACGATATATATATGAGAGAT	1269
Qy	2128	TATTCAGAGCATATCATTCATCATGTSAGGAATTAATATATCTCCAGAGAGCTGACAA	2187
Db	1270	CATACCAAGAGACGCTTCATCATGTGATGATTTCCCACTCCAGAGAACTGGCGA	1329
Qy	2188	GATATCGAAGGAAGTGCAGAAAACATATATATTTACCTAGTACTTATTAATGAGAGA	2247
Db	1330	TCACCTTATATCTTGACACCAAGAGAGAACAGTATTCACAGTGGAGGA	1389
Qy	2248	GCATTTCACTGTTAATCTTCCACAGATTTTGGGAATCACTGATCTGTTGGCTTGCATCA	2307
Db	1390	ACTTATGTTTCAAAAGAGTACATCTTGTGGCTTGAACATCTCTGTAAGGATTTGTATTA	1449
Qy	2308	TGTGAAGGACATCAAGATATATAGTCTGTGTGATTTAGAGAAAATGTTTGGAAATTA	2367
Db	1450	TCCTTAAAGGAATTAACACTATATAGCCGTTAGCGATCTGAAGGCGTTCGGGGTTA	1509
Qy	2368	A 2368	
Db	1510	A 1510	
RESULT 15			
AB035907			
LOCUS	AB035907	1092 bp	DNA linear
DEFINITION			xenopus laevis gene for xFTI, complete cds.
ACCESSION	AB035907		
VERSION	AB035907.1	GI:18146867	
KEYWORDS			xFTI.
SOURCE			xenopus laevis bone marrow DNA.
ORGANISM			Xenopus laevis
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
			Xenopodinae; xenopus.
REFERENCE			1 (sites)
AUTHORS			Kaneke,M., Nishihara,S., Kitano,T., Nariimatsu,H. and Saitou,N.
TITLE			The evolutionary history of glycosyltransferase genes
JOURNAL			unduplicated
REFERENCE			2 (bases 1 to 1092)
AUTHORS			Kaneke,M., Saitou,N. and Kitano,T.
TITLE			Direct Submission
JOURNAL			Submitted (17-DEC-1999) Mika Kaneke, National Institute of
			Genetics, Laboratory of Evolutionary Genetics, Yata 1111, Mishima,

Shizuoka 411-8540, Japan (E-mail: mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6785)

FEATURES
source location/Qualifiers

gene
1..1092 /organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_type="bone marrow"
1..1092 /gene="xft1"
1..1092 /gene="xft1"
/codon_start=1
/product="xft1"
/protein_id="BAB82490.1"
/db_xref="GI:18146868"
/translation="MRGSLLSVERLRRCRLCRSRRLRSRLAASLTCVLVYV
CSODLIAGINFLMPTLPKQVTLVLYRPLRPGKRLRCHLFLNFGCHITTR
SLYQADAIMPHHRDISPESDLPFRROSQKIMWNPESHSLSIGIFWMT
SYRVDLIFMRYGFLFSKRKATVLPKRKLVAWYISNNEDHERVQYINELNVEI
DYGRYGLDKEDNLVKTVESEKFLVLAENSLATDYITEKLIRNAFKSNALPIVMPGS
RYNMEFIPIRSSFIHVDDFSSPRKLAWYLKHLDKNIHLRYRYTWKKRYDVHTSFMD
EHYCTACOSVKAAGNHRASDLACWPFES"

BASE COUNT 256 a 281 c 294 g 261 t
ORIGIN

Query Match 7.9%; Score 222.6; DB 5; Length 1092;
Best Local Similarity 57.3%; Pred. No. 1.4e-29;
Matches 484; Conservative 0; Mismatches 349; Indels 12; Gaps 4;

OY 1472 AATGAACACTACTTCTGCGTGGGTGGCCATTGGGACACCTTTGACCTTACATCC 1531
DB 199 AAGCAAGTACGCTGCTTAATATGTATGAGCCATTGGCAACGGCAAGGCTTGGGAT 258
OY 1532 TGGCAAGCAATGTTCAACATCCAGAGTGCATCTCAACAGGACCGTTCACTGTACAC 1591
DB 259 TGGCAATGCTCTTCAATGTCAGGGGTGCGACATATCAACACAGAGCTTTTACAG 318
OY 1592 AAATCCATGCACTTCTGATCCATCACCGAGACATGATTGGATCTGACAAATTTACT 1651
DB 319 GAGCAGACGCCATTTATGTTCCACACCGGACAT--TTCTGATTTCACTGACCTTCT 375
OY 1652 CACCAAGCTAGGCGCCCTCCAGAAATGATTTGATGATTTGGAATACCACTCAC 1711
DB 376 TTCGAGAGGCGCCAGAGCTCCAAAGTGATCGATCGAATCTCCGCCAT 435
OY 1712 ACTCCCAAAAGAGTGGCATTTGACACTTGTTAACCTGACTGACTTACCGCGGAT 1771
DB 436 TCGCCCTGGCTGGCGAGTCTCGGGGCATCTTAACCTGACCATGTCTACAGGGTGAC 495
OY 1772 TCAGATATCCAAAGTGCCTTATGCTTCTTGACGCTAAGCACAATCCCTTCGTGTGAA 1831
DB 496 TCGATATCTTCAATGCGCTACGGCTTCT--ATTCTCAAGAAACGTGCAAAATAGTC 552
OY 1832 GTCCCAAGCAAGAAATTTGCTGCTGGGTGTGATGATGTAACCTTGACATGCC 1891
DB 553 CTGCGCGGCAAGAAAGCTGGTGGCTGTGATGAGCAACTGGAACGAGACAGAA 612
OY 1892 AGAGTCAAGTATACAGTACGCTAAGCAAAAGCATGAATCCATACCTACCGGCAAGCA 1951
DB 613 AGGTCACGATATTCACACGCTGAGGAACCTACGCAAAATCGACGTTTACGGCC--GG 669
OY 1952 TTTGAGAAATATGTCATGATAAAATTTGATTCCTACATATCTGTTGTAATTTAT 2011
DB 670 TATGGCTGGATTTAAAGAGCAACATCGTCAAGACCGTCTCGAGATATAATTTTAC 729
OY 2012 CTTTCCTTTGAAATTCATCCAGAGATTATACACGAAAGCTAT--ACAATGCT 2068
DB 730 TTGGCTTTTGAAGATTCCTGACAGGATATACACGAAAGCTGTGAGGAACGCG 789
OY 2069 TTTCTGGCTGGCTCTGATCTGTTTCCTGGGACATCTAGGAAACATATGAAATTTAT 2128
DB 790 TTTAAATCCAGCCATCCCATGTCATGAGGCCCAAGCCGCTATATACCTACGATGTTTC 849

OY 2129 ATTCCAGCAGATTCATTCATTCATGGAAGATTATTAATCTTCCAGTGAGCTAGCAAG 2188
DB 850 ATACCGCGAGCTCTTTCATTCACGTGAGCATTTCCAGGCCGAGGAAGCTGCCATG 909
OY 2189 TATCTGAGGAAGTGCACAAACAAATAGTTTATACCTAGTACTTTAATCGAGGAAG 2248
DB 910 TACCTGAGACCTGGATTAATAACATTCATCTGATACCGAGGATTTCACTGGAAGAG 969
OY 2249 GATTTCACGTAAATCTTCCACGATTTTGGGAATCACATGATGTTTGGCTTGCATCAT 2308
DB 970 CGTTACGAGGTGACGATGACGATCTTCTGAGACGAGCATTTACTGCAACCGGTCCAGAC 1029
OY 2309 GTGAA 2313
DB 1030 GTCAA 1034

Search completed: October 6, 2002, 04:30:09
Job time : 4349.27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 316.687 seconds
(without alignments)
15299.444 Million cell updates/sec

Title: US-09-744-748-4

Perfect score: 2822
Sequence: 1 aagcttgctactacagagag.....aacatgttcgatgaattc 2822

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq.032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2822	100.0	2822	21	AAZ92646
2	1542	54.6	2676	21	AAZ92647
3	1092.2	38.7	2036	21	AAZ92672
4	1092.2	38.7	2170	21	AAZ92645
5	127.4	4.5	1814	18	AAT59505
6	127	4.5	3594	18	AAT85219
7	114.4	4.1	1701	15	AAO77732
8	114.4	4.1	1701	21	AAZ21136
9	114.4	4.1	1701	21	AAA35014

10	114.4	4.1	6944	21	AAF21137
11	114.4	4.1	6944	21	AAA35015
12	113.4	4.0	1654	15	AAQ56905
13	113.4	4.0	1654	18	AAT61680
14	112.2	4.0	1086	15	AAQ56912
15	112.2	4.0	2042	18	AAT80111
16	112.2	4.0	2043	12	AAQ13300
17	112.2	4.0	2043	15	AAQ56906
18	112.2	4.0	2043	18	AAT76769
19	112.2	4.0	2043	18	AAT61675
20	111.8	4.0	1126	21	AAF21135
21	111.8	4.0	1126	21	AAA35013
22	111.8	4.0	1316	15	AAQ56911
23	111.8	4.0	1316	18	AAT61679
24	106.2	3.8	347	14	AAQ59752
25	103.6	3.7	795	17	AAT13798
26	102.4	3.6	2175	11	AAQ06691
27	100.8	3.6	1256	21	AAF21134
28	100.8	3.6	1400	13	AAA35012
29	100.8	3.6	1400	21	AAQ31436
30	100.8	3.6	1488	15	AAQ56910
31	100.8	3.6	2175	12	AAQ14382
32	100.8	3.6	2175	24	AAST1082
33	100.8	3.6	2861	12	AAQ14383
34	100.8	3.6	2861	21	AAF21133
35	100.8	3.6	2861	21	AAA35011
36	100.8	3.6	2861	24	AAST1083
37	100.8	3.6	3647	12	AAQ13333
38	100.8	3.6	3647	15	AAQ56909
39	100.8	3.6	3647	18	AAT61678
40	97.6	3.5	2134	18	AAT59506
41	73	2.6	8033	22	AAZ46308
42	69.2	2.5	6292	22	AAZ61117
43	67	2.4	5986	24	AAZ46735
44	66.6	2.4	1578	23	AAZ61433
45					ABL11867

ALIGNMENTS

RESULT 1	
AAZ92646	
ID	AAZ92646 standard; cDNA; 2822 BP.
XX	
AC	AAZ92646;
XX	
DT	05-JUN-2000 (first entry)
XX	
DE	Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:4.
XX	
KW	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW	brain; kidney; recombinant expression; transgenic animal; knockout
KW	animal; FDC-RIV; drug screening; inhibitor; potentiators; diagnosis;
KW	treatment; cancer; human; ss.
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
XX	/*tag= a
XX	/product= "Human alpha-1,3-fucosyltransferase"
PN	WO200006708-A1.
XX	
PD	10-FEB-2000.
XX	
PF	29-JUL-1999; 99WO-JP04092.
XX	
PR	29-JUL-1998; 98JP-0213823.
XX	
PA	(KIOW) KIOWA HAKKO KOGYO KK.
XX	

Human low adenosin
Human adenosine re
PCDNL1-alpha-(1-3)
Human alpha(1,3)-f
Lewis blood group
Fucosyltransferase
Glycosyltransferas
DNA encoding a gly
Human alpha 1,3/4
Human alpha(1,3/1,
Human low adenosin
Human adenosine re
DNA encoding a gly
Human alpha(1,3)-f
Human brain expres
Alpha-(1,3/1/4)-fu
CDX, a MTLA Invol
Human low adenosin
Human adenosine re
Encodes a Hela cel
DNA encoding a gly
Clone 7.2 encoding
Human cDNA clone 7
Clone 1 encoding 1
Human low adenosin
Human adenosine re
Human cDNA clone 1
GDP-Fuc:beta-D-Gal
DNA encoding a gly
Human alpha(1,3)-f
Human myeloid deri
Tumour suppressor
Human gene regulat
Human gene regulat
Drosophila melanog

PI Narimatsu H, Kudo T, Sasaki K;
XX WPI: 2000-183120/16.
DR P-PSDB: AAY80996.
XX
XX Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer.
XX
XX Claim 4: Page 134-142; 172pp; Japanese.

CC The invention relates to a novel alpha-1,3-fucosyltransferase which
CC transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglicosamine
CC (generating the Lewis x or y epitope). It does not transfer a fucose
CC moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglicosamine
CC and therefore does not generate the sialyl-Lewis x epitope. The
CC invention also relates to DNA sequences encoding alpha-1,3-
CC fucosyltransferase and expression vectors and host cells comprising
CC these DNA sequences. The invention additionally encompasses the
CC preparation of alpha-1,3-fucosyltransferase via the culture of
CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
CC screening potential inhibitors or potentiators of
CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
CC compounds having fucose-containing sugar chains by use of the protein;
CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
CC FUC-TV and is expressed mainly in brain and kidney tissues.
CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
CC potentiators and inhibitors may be used for the treatment and diagnosis
CC of diseases of the brain and kidney, and of cancers. They may be used for
CC the identification of substances which affect the activity or expression
CC of alpha-1,3-fucosyltransferase; such substances may be used
CC therapeutically. The knockout animals can be used to study the mechanisms
CC of action and expression of alpha-1,3-fucosyltransferase. Sequences
CC AA232646 and AA232647 represent cDNAs encoding human
CC alpha-1,3-fucosyltransferase (AAY80996).
XX
XX Sequence 2822 BP; 918 A; 485 C; 477 G; 942 T; 0 other;

Query Match 100.0%; Score 2822; DB 21; Length 2822;
Best local similarity 100.0%; Pred. No. 0;
Matches 2822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCTTGCTACTCAGAGAGTCTCATATGCTTTAAGAAAAATACATAGCAAAAGAGT 60
DB 1 AACCTTGCTACTCAGAGAGTCTCATATGCTTTAAGAAAAATACATAGCAAAAGAGT 60
QY 61 GCTTTTAAATGTCACATATATTTAGTAAATGAAGATATTTCCCGTTT 120
DB 61 GCTTTTAAATGTCACATATATTTAGTAAATGAAGATATTTCCCGTTT 120
QY 121 TTGGAAGGAACATCTAATGTTTCTATGCTTCTGCAAAATTTCAAAATCTTTAA 180
DB 121 TTGGAAGGAACATCTAATGTTTCTATGCTTCTGCAAAATTTCAAAATCTTTAA 180
QY 181 TTAGCTGGAATTTTATAGGCATATTTATTTCTGTAAGAGAAGACATGAATTTCAATG 240
DB 181 TTAGCTGGAATTTTATAGGCATATTTATTTCTGTAAGAGAAGACATGAATTTCAATG 240
QY 241 ACCCGTGAATTTTATATTAATTAACAGTTTGTGCAAAATATTTTGTGTCATAGCCAT 300
DB 241 ACCCGTGAATTTTATATTAATTAACAGTTTGTGCAAAATATTTTGTGTCATAGCCAT 300
QY 301 AAGAGGCGATGATGCGGAGAGAAAAAAGTGAAGAACTTTTCTTTAAAAAANA 360
DB 301 AAGAGGCGATGATGCGGAGAGAAAAAAGTGAAGAACTTTTCTTTAAAAAANA 360
QY 361 GCCTTCATCATCTTTTATAGCCACATCTGCTAAATTTCTGATGACATCAAGTGTCTTC 420
DB 361 GCCTTCATCATCTTTTATAGCCACATCTGCTAAATTTCTGATGACATCAAGTGTCTTC 420
QY 421 ATGAATATATAGCCCTTTTGTGGAATTTGTTTATTAATACGATATCATCTTAAT 480

DB 421 ATGAATATATAGCCCTTTTGTGGAATTTGTTTATTAATACGATATCATCTTAAT 480
QY 481 ACAAGAAAAAAGCTTTTAAAGCCCTTAGCTATGTAGAGAGAGTTCCTCAATGCA 540
DB 481 ACAAGAAAAAAGCTTTTAAAGCCCTTAGCTATGTAGAGAGAGTTCCTCAATGCA 540
QY 541 AATAGAACATTTGATTTATTAATACATCTGCTAGATATCGAGATGCTCAACAAATTTAANA 600
DB 541 AATAGAACATTTGATTTATTAATACATCTGCTAGATATCGAGATGCTCAACAAATTTAANA 600
QY 601 GTAAACAAAAAAGCTGTAGAAAAAATATTAATTAATAAACAACAGCTGTGAGTTTGA 660
DB 601 GTAAACAAAAAAGCTGTAGAAAAAATATTAATAAATAAACAACAGCTGTGAGTTTGA 660
QY 661 ATTTATTAATTTTGGCTGAGTGTACTGTTTAAAAAATGTATCAATCAAGAGTTGCC 720
DB 661 ATTTATTAATTTTGGCTGAGTGTACTGTTTAAAAAATGTATCAATCAAGAGTTGCC 720
QY 721 ATTTTATTTTCCAGATGATTTATATATATGATGCAAACTGTGAGGTAAATTTATACG 780
DB 721 ATTTTATTTTCCAGATGATTTATATATATGATGCAAACTGTGAGGTAAATTTATACG 780
QY 781 AGAAAAACAAAAACATATACGTTTAAACAGTTTGTGCTTACAGAAATTCATCTGA 840
DB 781 AGAAAAACAAAAACATATACGTTTAAACAGTTTGTGCTTACAGAAATTCATCTGA 840
QY 841 ACCACAAAGTTTGGTACAGTCTCAATATGCGAGTAATTAAGATTTATATCACTGGGA 900
DB 841 ACCACAAAGTTTGGTACAGTCTCAATATGCGAGTAATTAAGATTTATATCACTGGGA 900
QY 901 GATGTTACAAATCTGATGTGTGTTACAAAAATCTAAAAATTTAGTATGCTTTAGCGA 960
DB 901 GATGTTACAAATCTGATGTGTGTTACAAAAATCTAAAAATTTAGTATGCTTTAGCGA 960
QY 961 TGACAGTCTAGGTTGTCACACTTTTCTCAAGCAAAATTAAGAGAGTAAAGATAG 1020
DB 961 TGACAGTCTAGGTTGTCACACTTTTCTCAAGCAAAATTAAGAGATAG 1020
QY 1021 CATTTGAAAAAAGCTTACTGTTTGAAGATATCTGCTTTTGAAGAGTTTGGG 1080
DB 1021 CATTTGAAAAAAGCTTACTGTTTGAAGATATCTGCTTTTGAAGAGTTTGGG 1080
QY 1081 AAATCTCTTTTGAAGAGCTGTTCTGCTTTGAGGAAAGTCTGATCTCAATTTGAT 1140
DB 1081 AAATCTCTTTTGAAGAGCTGTTCTGCTTTGAGGAAAGTCTGATCTCAATTTGAT 1140
QY 1141 TGATTAACCTGAAAAAGAAATGCTGTGATTAATTAATGATTTATATATATGAT 1200
DB 1141 TGATTAACCTGAAAAAGAAATGCTGTGATTAATTAATGATTTATATATATGAT 1200
QY 1201 TTAAATATATCTCAATATATTTATGATTTCTCTCAATCCACGCTACCTCCCTTC 1260
DB 1201 TTAAATATATCTCAATATATTTATGATTTCTCTCAATCCACGCTACCTCCCTTC 1260
QY 1261 TGTCTTCTCTATTTTCTAGAGAAAAATATGACATCAACATCCAAAGAAATTTCTGCC 1320
DB 1261 TGTCTTCTCTATTTTCTAGAGAAAAATATGACATCAACATCCAAAGAAATTTCTGCC 1320
QY 1321 ATTTTAAATTTGCTGATATATCTGCTGCTGTTTATGATGCAATGCTTTTACATCA 1380
DB 1321 ATTTTAAATTTGCTGATATATCTGCTGCTGTTTATGATGCAATGCTTTTACATCA 1380
QY 1381 ACCTTACCAACAGCTGATCTTCACTCAATGCAATGCAAGCAAGCTGTGCTGTAAGTGA 1440
DB 1381 ACCTTACCAACAGCTGATCTTCACTCAATGCAATGCAAGCAAGCTGTGCTGTAAGTGA 1440
QY 1441 AAATCTCTTTTCCACCAAAAGATATTTAAAGAACTACTATCTGCTGCTGCTG 1500
DB 1441 AAATCTCTTTTCCACCAAAAGATATTTAAAGAACTACTATCTGCTGCTGCTG 1500
QY 1501 GCCATTTGGGAGACCTTTGACCTTACATCTGCTCAAGCAAGTGTGAATCAACGAAGATG 1560
DB 1501 GCCATTTGGGAGACCTTTGACCTTACATCTGCTCAAGCAAGTGTGAATCAACGAAGATG 1560

```

Db 1501 GCCATTGGGACAGACCTTTGACCTTACATCTGCCAAGCAATGTTCAATCCAGAGATG 1560
Oy 1561 CCATCTCACAAGSAGACCTTTCATCTGACAAACAATCCATGAGATTTGATCATCACC 1620
    |||||
Db 1561 CCATCTCACAAGSAGACCTTTCATCTGACAAACAATCCATGAGATTTGATCATCACC 1620
Oy 1621 AGACATCACTGGGATCTGACAAATTTACCTCAGCAAGCTAGGCCACCTTCCAGAAATG 1680
    |||||
Db 1621 AGACATCACTGGGATCTGACAAATTTACCTCAGCAAGCTAGGCCACCTTCCAGAAATG 1680
Oy 1681 GATTGGAGTAATTTGGATTCACCAACTCCCACTCCCAAAAGAGGATGAGACACT 1740
    |||||
Db 1681 GATTGGAGTAATTTGGATTCACCAACTCCCACTCCCAAAAGAGGATGAGACACT 1740
Oy 1741 GTTTAACTGACTGACTTACCGCGGATTCAGATATCCAAAGTCCCTATGAGCTTCT 1800
    |||||
Db 1741 GTTTAACTGACTGACTTACCGCGGATTCAGATATCCAAAGTCCCTATGAGCTTCT 1800
Oy 1801 GACGGTAAGCACAATCCCTTGTGTGTAAGTCCCAAGCAAAAGAGAAATGGTGTGCTG 1860
    |||||
Db 1801 GACGGTAAGCACAATCCCTTGTGTGTAAGTCCCAAGCAAAAGAGAAATGGTGTGCTG 1860
Oy 1861 GGTGTGAGTAAGTGAACCTGAGCATGCCAGAGTCAAGTATTAATGAGCTAGACAA 1920
    |||||
Db 1861 GGTGTGAGTAAGTGAACCTGAGCATGCCAGAGTCAAGTATTAATGAGCTAGACAA 1920
Oy 1921 AAGCATTTGAATTCATACCTAGCGGCAAGCATTTGGAGATATGTCATGATATAAATTT 1980
    |||||
Db 1921 AAGCATTTGAATTCATACCTAGCGGCAAGCATTTGGAGATATGTCATGATATAAATTT 1980
Oy 1981 GATTCCTACCAATATCTGCTTGTAAATTTTATCTTCCCTTTGAAATTTCAATCCACAAGA 2040
    |||||
Db 1981 GATTCCTACCAATATCTGCTTGTAAATTTTATCTTCCCTTTGAAATTTCAATCCACAAGA 2040
Oy 2041 TTACATTCAGSAGAAAGCTTATACATGCTTTCTGGCTGCTCTGTAACCTGCTGTCTGG 2100
    |||||
Db 2041 TTACATTCAGSAGAAAGCTTATACATGCTTTCTGGCTGCTCTGTAACCTGCTGTCTGG 2100
Oy 2101 ACCATCTAGSAGAAAGCTATGAGATTTATTTCCAGAGATTCATTCATGATGAGAA 2160
    |||||
Db 2101 ACCATCTAGSAGAAAGCTATGAGATTTATTTCCAGAGATTCATTCATGATGAGAA 2160
Oy 2161 TTATTAATCTCCAGTACCTAGCAAAAGTATCTGAAGAGTGCAGCAAAACAATTAAGTT 2220
    |||||
Db 2161 TTATTAATCTCCAGTACCTAGCAAAAGTATCTGAAGAGTGCAGCAAAACAATTAAGTT 2220
Oy 2221 ATACCTAGTACTTAACTGAGGAGGATTTCACTGTAATCTTCCAGATTTTGGGA 2280
    |||||
Db 2221 ATACCTAGTACTTAACTGAGGAGGATTTCACTGTAATCTTCCAGATTTTGGGA 2280
Oy 2281 ATACATGATGATTTGGCTTGGCATGATGAAAGGATCAAGATATATAGTCTGTGG 2340
    |||||
Db 2281 ATACATGATGATTTGGCTTGGCATGATGAAAGGATCAAGATATATAGTCTGTGG 2340
Oy 2341 TTAATTTAGAGAAATGGTTTGAATTTAAATTTTCATCACTTGCACACTGATATAAT 2400
    |||||
Db 2341 TTAATTTAGAGAAATGGTTTGAATTTAAATTTTCATCACTTGCACACTGATATAAT 2400
Oy 2401 TTTGATGAGATTCATCCAGATGAGATTAAGAAAGATGCAATCTCTCTTTTGG 2460
    |||||
Db 2401 TTTGATGAGATTCATCCAGATGAGATTAAGAAAGATGCAATCTCTCTTTTGG 2460
Oy 2461 TCACATATTTATTTATCACCTCTCTAGGATGAGTATATTTTGGAGATTTTGA 2520
    |||||
Db 2461 TCACATATTTATTTATCACCTCTCTAGGATGAGTATATTTTGGAGATTTTGA 2520
Oy 2521 AAAGCTCAGCATGAGCAATTCATTCGTTTAAATTTATTCGATATATACCTAATTA 2580
    |||||
Db 2521 AAAGCTCAGCATGAGCAATTCATTCGTTTAAATTTATTCGATATATACCTAATTA 2580
Oy 2581 TGTGACAGTGAAGTAATTTATTTATTCATTTGTAACCAATTTGCTTTTACATTT 2640
    |||||
Db 2581 TGTGACAGTGAAGTAATTTATTTATTCATTTGTAACCAATTTGCTTTTACATTT 2640

```

```

Oy 2641 TTGATGTCATATATGTAAGCTTGTGGTTGATTAATTTTCCACACTGATACCTGT 2700
    |||||
Db 2641 TTGATGTCATATATGTAAGCTTGTGGTTGATTAATTTTCCACACTGATACCTGT 2700
Oy 2701 TTAATCTATTTGGGAATAGATGACATCTTAAAGTATGAAAAATTTTCACTAATAT 2760
    |||||
Db 2701 TTAATCTATTTGGGAATAGATGACATCTTAAAGTATGAAAAATTTTCACTAATAT 2760
Oy 2761 TCAATGCTAGTCCAACTTTCATATATTAACAGAGAAACATGTTGGATGAA 2820
    |||||
Db 2761 TCAATGCTAGTCCAACTTTCATATTAACAGAGAAACATGTTGGATGAA 2820
Oy 2821 TC 2822
    ||
Db 2821 TC 2822

RESULT 2
AAZ92647
ID AAZ92647 standard; cDNA; 2676 BP.
XX
AC AAZ92647;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.
XX
KW Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW brain; kidney; recombinant expression; transgenic animal; knockout
KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
KW treatment; cancer; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 295..1374
FT /*tag= a
FT /product= "Human alpha-1,3-fucosyltransferase"
XX
PN WO200006708-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-JP04092.
XX
PR 29-JUL-1998; 98JP-0213823.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Narimatsu H, Kudo T, Sasaki K;
XX
DR WPI: 2000-183120/16.
DR P-PSDB: AAT80996.
XX
PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer.
XX
PS Claim 4; Page 143-150; 172pp; Japanese.
XX
CC The invention relates to a novel alpha-1,3-fucosyltransferase which
CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
CC (generating the Lewis x or y epitope). It does not transfer a fucose
CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
CC and therefore does not generate the sialyl-Lewis x epitope. The
CC invention also relates to DNA sequences encoding alpha-1,3-
CC fucosyltransferase and expression vectors and host cells comprising
CC these DNA sequences. The invention additionally encompasses the
CC preparation of alpha-1,3-fucosyltransferase via the culture of
CC transformed cells or by expression of the protein in a transgenic animal;
CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
CC screening potential inhibitors or potentiators of

```



```

Db 947 AGAATTAATATCCAGCTGATTCATTCATCATGGAAGATTTAACTCCACAGTAGT 1006
QY 2181 TAGCAAGATATCTGAAGAGATGACACAAAACAATAAGTATACCTAGTACTTAACT 2240
Db 1007 TAGCAAAATATCTGAAGAGATGACACAAAACAATAAGTATACCTAGTACTTAACT 1066
QY 2241 GGAAGAAAGATTTCACTGTAATCTTCACAGATTTGGGAATCAATGATGATTTGGCTT 2300
Db 1067 GGAAGAAAGATTTCACTGTAATCTTCACAGATTTGGGAATCAATGATGATTTGGCTT 1126
QY 2301 GGCATCATGGAAGAAAGATGACACAAAACAATAAGTATACCTAGTACTTAACT 2360
Db 1127 GGCATCATGGAAGAAAGATGACACAAAACAATAAGTATACCTAGTACTTAACT 1186
QY 2361 GGAATTAATATCCAGCTGATTCATTCATCATGGAAGATTTAACTCCACAGTAGT 2420
Db 1187 GGAATTAATATCCAGCTGATTCATTCATCATGGAAGATTTAACTCCACAGTAGT 1242
QY 2421 GATTTGAGATTAAGAAAGATGACACATACTACTTTGTCACAAATTTATTTATCAC 2480
Db 1243 GATTTGAGATTAAGAAAGATGACACATACTACTTTGTCACAAATTTATTTATCAC 1297
QY 2481 CCTCTAGAGGTACAGTATATTTTGGTGAGATTTTAAAGCTCAGCATGACATC 2540
Db 1298 CCTCTAGAGGTACAGTATATTTTGGTGAA-ATTTTAAAGATCAGATTAAGCATC 1356
QY 2541 ATTCATCGTGGTTTAAATATCTGATATACCTATATTTGACACTGAGATATTT 2600
Db 1357 ATTCATCGTGGTTTAAATATCTGATATATCTGATATATGATACGACACTGGA-ATAATATTT 1415
QY 2601 ATTCATCGTGGTTTAAATATCTGATATATCTGATATATGATACGACACTGGA 2660
Db 1416 ATTCATCGTGGTTTAAATATCTGATATATCTGATATATGATACGACACTGGA-ATAATATTT 1472
QY 2661 AGTTGTGGTTTATTTTTCACACTGATGATGTTTATATCTATTTGGGAATGA 2720
Db 1473 AGTTGTGGTTTATTTTTCACACTGATGATGTTTATATCTATTTGGGAATGA 1529
QY 2721 AGATGACATCTTAAGTATGA-AAAATTTTCACTAAGTATTAACAATGCTAGTCCAAC 2779
Db 1530 AGATGACATCTTAAGTATGA-AAAATTTTCACTAAGTATTAACAATGCTAGTCCAAC 1589
QY 2780 TTTGATCTATTAACAGAGAGAGAA 2804
Db 1590 TTTGATCTATTAACAGAGAGAGAA 1614

RESULT 5
AAT59505
ID AAT59505 standard; cDNA; 1814 BP.
XX
AC AAT59505;
XX
DT 06-MAY-1997 (first entry)
XX
DE Murine myeloid-lineage alpha-(1,3)-fucosyltransferase cDNA.
XX
KW Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM;
XX septic shock; septicemia; therapy; ss.
XX
OS Mus sp.
XX
FH key Location/Qualifiers
FT CDS 325..1353
FT /tag= a
XX
PN W09640881-A1.
XX
PD 19-DEC-1996.
XX
PF 08-MAY-1996; 96MO-US06427.
XX

```

```

PR 07-JUN-1995; 95US-0483151.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Holgersson J, Seed B;
XX
DR WPI; 1997-108639/10.
XX
DR P-PSDB; AAW11820.
XX
PT New murine alpha-(1,3)-fucosyltransferase - for fucosylating an
XX antibody to protect mammals against e.g. septic shock or septicemia
XX
PS Claim 6; Fig 3; 58pp; English.
XX
CC A cDNA clone (AAT59505) codes for murine myeloid lineage alpha-
CC (1,3)-fucosyltransferase (AAW11820), an enzyme which has a relatively
CC strict substrate requirement for sialylated N-acetyllactosamine,
CC which can account for the presence of the sialyl-Le-w epitope on
CC murine cells, and which is more effective than Fuc-TIV in support
CC of E-selectin-mediated COS cell adhesion. It was isolated from a
CC murine myeloid 32D cl3 cell cDNA library by identification of a
CC clone capable of directing the expression of sialyl-Le-x
CC determinants. Transformed host cells (e.g. 32D cl3 or human 293
CC cells) expressing the alpha-(1,3)-fucosyltransferase, and pre-
CC another fucosyltransferase such as human Fuc-TIV (see also AAT59506),
CC can be used to fucosylate an antibody (e.g. IgG or IgM) for use in
CC protecting an animal against an adverse immune reaction, esp. septic
CC shock or septicemia.
XX
SQ Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other;

Query Match 4.5%; Score 127.4; DB 18; Length 1814;
Best Local Similarity 52.1%; Pred. No. 8e-16;
Matches 334; Conservative 0; Mismatches 301; Indels 6; Gaps 2;

QY 1557 GATGCCATCTCACAGGACCGTTCAGTACACAAATCCATGACTTGTATTCATC 1616
Db 548 GCTGCCCTCTGAGTGTCAACCGGAGCTGCTAGCAGTGTGCTGTCCACC 607
QY 1617 ACCGAGCATCACTGGGATCTGACAAATTTACTCTGACAGGACGACCCCTTCAGA 1676
Db 608 ACCGAGCTGACAAACCGGCAATCTCTCCATCCCTGGACCAAGGACACAGGAGAC 667
QY 1677 AATGATTTGATGAATTTGGAATCACCACTACATCCCAAAAGAGGATTTAGC 1736
Db 668 CTGGGCTGCGGCTCCATGGAATCGCCAGTAAATACCATGCTCTCCATCCGCTG 727
QY 1737 ACTTGTAACTGACTCTGACTTACCGCGGTGATTCAGATTCAGATTCAGTTCG 1796
Db 728 GCATCTTCAACTGGTCTGAGTATGCGGTGATTCAGATTCAGTATCCCTACGCTG 787
QY 1797 TCTTGACGTTAAGCACAATCCCTCTCTTTGAAGTCCCAACCAAAAGAAATTTG 1856
Db 788 GCTTGAGACCTCTCTCTGGGCC--CACATCCCACTACCGGCCCAAAAGAGATTTG 844
QY 1857 GCTGGGTTGTGAGTAAGTGAACCTGAGCATGACAGTCAAGTATTAATGAGCTAA 1916
Db 845 CTTGGGTGATCAGCAATTTCCAGAGAGGCGACAGCATGCAAGCTTACCGCAGCTGG 904
QY 1917 GCAAAAGCATTTGAATTCATCTACGGGCAAGCATTTGGAGAAATATGCAATGAAA 1976
Db 905 CCCCTCATCTGACAGTGAATGTGTGCTGCGCCAGCGAGGAGCCCTATGCGTAAT 964
QY 1977 ATTTGATCTTACCATATCGCTTGAATTTTATCTTCTTGAATAATTCATTCACA 2036
Db 965 GTCTGCTGCCCTTTGGCCGGGTACCGCTTTCATCTGAGCTTTGAGAACTCAAGCAT 1024
QY 2037 AGGATTACATCAGGAAAGCAAT--ACAATGCTTTTCTGCTGCTGCTGATCCTGTTG 2093
Db 1025 GGGACTACATCACTGAGAAATTTCTGGGCAATCCCTGGGGGCTGTGCTGATCCGCTG 1084
QY 2094 TTCTGGGACCATTTAGGAAACTATGAGAAATATATTCAGAGATTCATTCATTCAGT 2153

```


Tue Oct 8 10:18:24 2002

us-09-744-748-4.rng

Page 8

```
DB 1085 CGCTGGAGCCTCCTCGGGCCACCTACGAGCCTTTGTGCCACCAAGTGCCTTTGACACG 1144
OY 2154 TGGAGATTTAACTCTCCCACTGAGCTAGCAAAAGTATCTG 2194
DB 1145 TGGACGACTTCAGCTCTGCGCCGTGAACTGCTGTTCTTCCTG 1185

RESULT 6
AAT85219 standard; DNA; 3594 BP.
XX
AC AAT85219;
XX
DE 02-MAR-1998 (first entry)
XX
XX Mouse alpha-fucosyltransferase Fuc-TVIII gene.
XX
XX Alpha-fucosyltransferase; Fuc-TVII gene; mouse;
XX sialyl Lewis x determinant; oligosaccharide; antiinflammatory;
XX inhibitor; ds.
XX
OS Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 966..3082
XX FT /*tag= a
XX FT /note= "includes introns"
XX FT 465..477
XX FT /number= 1
XX FT 478..669
XX FT /*tag= c
XX FT /number= 1
XX FT 670..1149
XX FT /*tag= d
XX FT /number= 2
XX FT /codon_start= 966..998
XX FT 1150..1946
XX FT /*tag= e
XX FT /number= 2
XX FT 1947..1959
XX FT /*tag= f
XX FT /number= 3
XX FT /codon_start= 1947..1949
XX FT /note= "alternative ATG start codon"
XX FT 1960..2065
XX FT /*tag= g
XX FT /number= 3
XX FT 2067..3552
XX FT /*tag= f
XX FT /number= 4
XX FT /codon_start= 2126..2128
XX FT /note= "alternative ATG start codon"
XX
XX PN M09732889-A1.
XX
XX PD 12-SEP-1997.
XX
XX FE 07-MAR-1997; 97MO-US03184.
XX
XX PR 08-MAR-1996; 96US-0613098.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Gersten KM, Lowe JB, Natsuka S;
XX
XX PI MPI: 1997-470535/43.
XX
XX DR P-PSDB; AAM27138 and AAM26671.
XX
XX PT Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g.
XX screening for inhibitors used as antiinflammatory agents
XX
XX Claim 1; Fig 2; 91pp; English.
```

```
XX This DNA sequence comprises the mouse Fuc-TVIII gene encoding
CC alpha-fucosyltransferase. It was isolated from a 373 cell genomic
CC library by screening with a probe from the catalytic domain of the
CC human Lewis alpha(1,3/1,4)fucosyltransferase (Fuc-TVIII).
CC Examination of the mouse Fuc-TVIII locus identified 3 Met codons
CC that may initiate translation of alpha(1,3) fucosyltransferases
CC with different cytosolic domains encoded by exons 2 and/or 3, but
CC with identical Golgi-localised catalytic domains encoded by exon 4
CC (see AAM27138 and AAM26671). The proteins and nucleic acids can
CC be used to construct animal cell lines able to post-translationally
CC modify oligosaccharides on cell surface, intracellular or secreted
CC proteins and lipids, to isolate reagents for efficient enzymatic
CC production of oligosaccharides, to generate antibodies to
CC glycosyltransferases useful as diagnostic reagents, to screen for
CC fucosyltransferase inhibitors and inactivators, especially those
CC that act as antiinflammatory agents, for genotyping individuals at
CC the fucosyltransferase locus and for in vitro synthesis of sialyl
CC Lewis x tetrasaccharide.
XX
SQ Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other;
```

```
Query Match 4.5%; Score 127; DB 18; Length 3594;
Best Local Similarity 52.3%; Pred. No. 1,1e-15;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
```

```
OY 1557 GATGCCATCTCCACACGAGCCGTTCACTGATACAAACAAATCCATGCTGTGATCCATC 1616
DB 2277 GCTCCGCTCTGAGTGTGACAGGAGCTGCTAGCCAGTGTGTGTGGTCTTGCTTCACAC 2336
OY 1617 ACCGAGACATCATGTTGGAGTCTGACAAATTTACCTCAGCAAGTACGCCACCTTCAGAG 1676
DB 2337 ACCGTGAGCTGCAAAACCCGCAATCTCTCCATCCCTGACAGAGCCCAACAGAGACG 2396
OY 1677 AATGATTTGATGATTTTGAATCACCACACTCCACCTCCCAAAAGAGTGGCATTTAGC 1736
DB 2397 CTTGGGTCTGGGCTCATGATGATGCGCCAGATATATATATGATGATGATGATGATG 2456
OY 1737 ACTGTTTAACTGACCTGACCTTACCGCGTGATTCAGATATATATGATGATGATGATG 1796
DB 2457 GCATCTTCAACTGGGCTGAGCTATGCGCTGATTCAGATATATGATGATGATGATG 2516
OY 1797 TCTTGACGATAGCAACAATCCCTTGTGTTGATGCCAAGCAAGAAATTTGCTGT 1856
DB 2517 GCTTGAAGCCTCTCTCTGTGGCC--CACATCCCACTACCGGCAAAAGCATGCTGTG 2573
OY 1857 GCTGGGTTTGAGTAACTGGAAACCTTGAGCATGCCAGATGCAATTTACATGAGCTAA 1916
DB 2574 CTTGGGTGATCAGCATTTTCCAGAGCGCGAGCGTGCAGAAAGCTGACGCGAGCTGG 2633
OY 1917 GCAAAGCATTTGAATCCATCTACTTACCGGCGCAAGATTTGGAGAAATGTCAATGTAAA 1976
DB 2634 CCCCCTATCTGACAGTGGATATGTTGCTGCCGCCGACCGAGCGCCCTATGACCTATTT 2693
OY 1977 ATTGATTCCTACCATATATGCTTTGAATTTATCTTCTTGTGAATTTATCATCCACA 2036
DB 2694 GTCTGTGCTCCCTTGTGGCCGCTGACCGCTTACTGCGCTTTGAGAACTCAGACATTC 2753
OY 2037 AGGATTTACATCAGGAAAGCTAT--ACAATGCTTTTCTGCTGCGCTGTACCTGTG 2093
DB 2754 GGGACTCATCATCTAGAGAGTCTGCGCATGCGCTGGGGGTGGTGTGTGTACCCCTGG 2813
OY 2094 TTTGAGGACCATCTAGGAAACTATGATTTATTCAGAGATTCATCATTCATTCATG 2153
DB 2814 CGCTGGGACCTCTCTGGGCGCACTTCAGAGGCTTTTGTGGCACAAGATGCTTTGTACACG 2873
OY 2154 TGGAGATTTAACTCTCCCACTGAGCTAGC 2184
DB 2874 TGGACGACTTCAGCTCTGCCCTGAACTGGC 2904
```

RESULT 7
AA077732

[illegible]

QY	1904	TTCATGACCTAGACAAAAGCATTTGAATCCATACCTACGGGCAAGCATTTGGAGAAATAT	1963
Db	690	TACGGGACAGCTGGCGGCTTCATCTGGGGGATGTCCTTTGGCCGTCGCAATGGACGGCCA	749
QY	1964	GTCATGATGATAAAATTTGATTTCTTCACCATATCTGGCTTGTAATTTTATCTTCCCTTGAA	2023
Db	750	CTGTGGCCAGCTGCGCTGCTGGTGGCCACCGTGGCCCACTACCGCTTCTACCTGTCTTTGAG	809
QY	2024	AATTCGAATCCACAGAGATTCATCATACGGAAGAAAGCTAT--ACAAATGCTTTTCTGGCTGGC	2080
Db	810	AACTCTCAGACACCGGAGACTACATATACGGAGAAATCTGGCGGCAACGACACTGGTGGCTGGC	869
QY	2081	TCTGTAACCTGTTGTTCTGTGGGACCATCTAGGGAACAACTATGAGATTAATATTCACAGAGAT	2140
Db	870	ACTGTGCCAGTGTGCTGGGGGCCCCACGGCCCACTATGAGAGCCCTTGCTGGCCGGCTGAC	929
QY	2141	TGATTCATTCATGATGTAAGATTAATATTAACCTGCCAGTGAGCTAGC	2184
Db	930	GCGTTCGTGCATGTGATGATGACTTTGGCTCAGCCGACGAGACTGGC	973
RESULT 8			
AAE21136			
ID	AAE21136 standard; DNA; 1701 BP.		
AC	AAE21136;		
XX			
XX			
DT	14-MAR-2001 (first entry)		
DE	Human low adenosine antisense oligonucleotide related sequence #2703.		
XX			
XX	Low adenosine antisense oligonucleotide; phosphorothioate; allergy;		
KM	human; airway disorder; bronchococci; lung inflammation;		
KM	surfactant depletion; respiratory; bronchodilator; antiinflammatory;		
KM	immunosuppressive; antiallergic; hypotensive; cytostatic;		
KM	respiratory obstruction; pulmonary obstruction; impeded respiration;		
KM	surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;		
KM	respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis		
KM	pulmonary hypertension; emphysema; pulmonary transplantation rejection;		
KM	chronic obstructive pulmonary disease; pulmonary infection; bronchitis;		
KM	cancer; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO20062736-A2.		
XX			
PD	26-OCT-2000.		
XX			
PE	24-MAR-2000; 2000WO-US08020.		
XX			
PR	06-APR-1999; 99US-0127958.		
XX			
PA	(UYEC-) UNIV EAST CAROLINA.		
XX	(NYCE/) NYCE J W.		
XX			
PI	Nyce JW;		
XX			
DR	WPI; 2000-679539/66.		
XX			
PT	Low adenosine (A) content antisense oligonucleotides which do not		
PT	trigger adenosine receptors during metabolism, useful e.g. for treating		
PT	cancers and respiratory obstructions -		
XX			
PS	Disclosure; Page 1034; 1592pp; English.		
XX			
CC	The present invention describes low adenosine (A) content antisense		
CC	oligonucleotides and compositions (I) comprising them. In the antisense		
CC	oligonucleotides the A is replaced by a 'universal' or alternative base		
CC	(I) can have respiratory, bronchodilator, antiinflammatory, analgesic,		
CC	immunosuppressive, antiallergic, hypotensive and cytostatic activities		
CC	The antisense oligonucleotides and (I) can be used to down-regulate the		
CC	expression and or activity of target polypeptides associated with		

CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergies
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impaired respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.

XX Sequence 1701 BP; 290 A; 527 C; 589 G; 295 T; 0 other;

Query Match 4.1%; Score 114.4; DB 21; Length 1701;

Best Local Similarity 50.8%; Pred No. 2.8e-13;
 Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

QY 1544 TTCAACATCCAGAGATCCATCAGACGACCGTTCAGTACAAACAAATCCCATGCA 1603
 DB 333 TACGGCATCGCCCGCTCCACCTGAGTGCACCAACGACCTGCGGCGGAGCGCC 392
 QY 1604 GTTCATCATCACCAGACATCAAGTGGATCTGACAAATTTACTCAGCAAGTAGG 1663
 DB 393 GTGGCTCTCCACACCGCGAGCTGACACCGCGGTCCACCTCCGCGCCACCGG 452
 QY 1664 CCACCGCTCCAGAAATGATTTGATGAAATTTGAAATACCAACACACGCGCCAAAAG 1723
 DB 453 CCGGAGGGGACGCGCTGGGTGGCTCATGAGTCTCTAGCAGACACCGAGGCTC 512
 QY 1724 AGTGGCATTTGAGACATTTGATCTGATCTGACTTACCTCCGCGGATTCAGATTC 1783
 DB 513 ACCCATCTCGAGCATCTTCACTGAGTGGTCTGAGCTTCCGCGGATTCGAGATCTTT 572
 QY 1784 GTGCTTATGCTCTTCTTACAGTAAACAAATCCCTTGTGTTGAAGTCCAAAGCAA 1843
 DB 573 GTGGCTTATGAGCGCGCTGGGTGGGAGCGCCACCTGGGGGCGCT--CGCCACCGCTCCGAGCCAG 629
 QY 1844 GAGAAATTTGTCTGCTGGTGTGATTAACAGTAAACCTGAGCAGTCCAGATCAATAT 1903
 DB 629 AGCAGAGTGGCGCGCTGGGTGGTGCACAACTTCCAGAGCGGAGCTCGTGGGAGCTG 689
 QY 1904 TCAATGAGCTAAGCAAAACATTTGAATTCATACCTAAGCGGCAAGCTTTGGAGATAT 1963
 DB 690 TACCGGAGCTGGGGGCTCATCTGCGGGGTGATGATCTTTCGCGCAAGAGCGCCA 749
 QY 1964 GTCAATGATATAAATTTGATCTTACCATATCTGCTTGAATTTTATCTTCTTGAA 2023
 DB 750 CTGTGGCGGACGCTGGGTGGGCGGAGCGGAGTCCGCTTCTTACCTGCTCTTGGAG 809
 QY 2024 AATTCATTCACCAAGATTAATCAGGAAAGCTAT--ACAATGCTTTTCTGGCTGGC 2080
 DB 810 AACTCTGACGACCGGCTCATTAAGGAAATTTCTGGCCACAGCAGCTGGTGGCGG 869
 QY 2081 TCTGTACCTGTTTCTGGGAGCATTAAGGAAATTTGAGATTTATTTATCTCAGAGAT 2140
 DB 870 ACTGTGCGAGTGTGTGGGGGCGGAGCGGCGCACTTATGAGGCTTCTGTGGCGGCTGAC 929
 QY 2141 TCATTCATTCATTCGAGAAATTAATCTCCAGTAGAGCTAGC 2184
 DB 930 GCCTTGTGATGATGATGATTTGCTGACCGGAGAGCTGGC 973

RESULT 9
 AAA35014
 ID AAA35014 standard; DNA: 1701 BP.

AAA35014;

28-JUL-2000 (first entry)

Human adenosine receptor related polynucleotide SEQ ID NO:2703.

KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiaesthetic; cyostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

OS Homo sapiens.

PN MO200009525-A2.

PD 24-FEB-2000.

PF 03-ADG-1999; 99WO-US17712.

PR 03-ADG-1998; 98US-0095212.

(UYEC-) UNIV EAST CAROLINA.

Nyce JW;

WPI; 2000-205971/18.

PT New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers -

PS Disclosure: Page 957-958; 1343pp; English.

CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cyostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing the
 CC bronchoconstriction and inflammation. AAA32313 to AAA3512 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 2815, and then the last
 CC differ from the previously named sequences. SEQ ID NO:11 to 1880
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of ONs from the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.

XX Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;

Query Match

Best Local Similarity 4.1%; Score 114.4; DB 21; Length 1701;
 50.8%; Pred. No. 2.8e-13;

AA056905	
ID	AA056905 standard; DNA; 1654 BP.
XX	
AC	AA056905;
XX	
DT	26-JUL-1994 (first entry)
XX	
DE	pCDNA1-alpha-(1-3)Fuc-TyI genomic DNA.
XX	
KM	Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
KW	surface; oligosaccharide; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CDS	128..1208
FT	/tag= a
XX	
PN	W09402616-A.
XX	
PD	03-FEB-1994.
XX	
PE	20-JUL-1993; 93WO-US06703.
XX	
PR	20-JUL-1992; 92US-0914281.
XX	
PA	(UNMI) UNIV MICHIGAN.
PI	Lowe JB;
DR	WPI: 1994-048874/06.
DR	P-PSDB: AAR45933.
XX	
PT	DNA fragment encoding a glycosyltransferase - can be used for in
PT	viro reactions to modify cell surface oligosaccharide(s) e.g.
PT	blood gp. determinants, to protect against transplant rejection
XX	
PS	Claim 3; Fig 7; 249pp; English.
XX	
CC	The sequence is that encoding human GDP-Fuc: [beta-D-Gal (1-4)]-D-
CC	Glucose alpha (1-3)-fucosyltransferase. The enzyme produced by the
CC	DNA may be non-glycosylated. This prevents premature loss of enzyme
CC	activity. It can also be used in in vitro reactions to modify cell
CC	surface oligosaccharide mois. e.g. blood group determinants.
CC	See also AA056906-12.
XX	
SQ	Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;
<hr/>	
Query Match	4.0%; Score 113.4; DB 15; Length 1654;
Best Local Similarity	51.2%; Pred. No. 4.4e-13;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps	5
OY	1484 ATTCTGCTGTGGTGTGGCATTTTGCGCAGACCTTTGACTTCATCTGCCAAGCAATG 1543
Db	318 ATTCGCTGTGCACCTGGCTTTTAACAACCATTAGCTCTCCCCGGCTCAGAGATG 377
OY	1544 TT---CAACATCCAAGATGCCATCTCACCAAGGACCGTTCACTGNACAACAATGCCAT 1600
Db	378 GTGCTGGCAGCGCTGACTGCTCAACATCACTGCCGACCCGAAGTGTATCCACAGGCAGC 437
OY	1601 GCAGTTTGATTCATCCACCGAGACATCAAGTTGGATCTGA---CAAATTTACCTGACAA 1657
Db	438 GCGGTCATCGTCCACACCGAGAGTCATGTACAAACCCCAAGGCCOAGCTCCACGCTCC 497
OY	1658 GCTAAGCCACCTTCACAGAATGATTTTGATGTAATTGGAAATCACCAACTCACACTCCC 1717
Db	498 CGGAGCGGCAAGGGCAGGATGATCTGGTTCAACATGAGATGCCCAAGCCACTGCTGG 557
OY	1718 CAAAAGAGTGGATTCAGACATCTGTTTAACTGCATCTGACCTTAACCCCGTGTATTCAGAT 1777
Db	558 CAGCTGAAAAAGCCATGAGCGGATTACTTCAATCTCACCATATCTCTACCGCAGCGATCCGAC 617
OY	1778 ATCCAAAGTGCCTTATAGCT-----TCTTAGAGGTAAAGCAAAATCCCTTCGTGTTT 1828

Db	618	ATCTTCACGCCCCTACGGCTGGCTGGAGCCGTGGTCCGGCCAGCCTTGCACCCACCCTC	677
Qy	1829	GAAGTGGCCAAACCAAGAGAAATGGTGTGCTGGGTTGTGAGTAACCTGAGCAAT	1888
Db	678	AACTCTCGGCGCAAGACCGAGCTGTGTGGCCTGGGGCAGTGTCCAACTGGGGGCCAATCTCC	737
Qy	1889	GCCAGAGTCAAGTATTTCATTAAGCTTAGCTTAGCAAAAGCATTGAAATCCATCTACGGGCNA	1948
Db	738	GCCAGGGGTGGCTACTCTCCAGAGCCTGCAGGCCCATCTCAAGGTGGACGTGTACGG---A	794
Qy	1949	GCATTGTGAGAAATATGTCATATGATTAATAAATTTGATTCATACATATCTCTGTAAATTT	2008
Db	795	CGCTCCCAACAGCCCCCTGGCCCCAGGGACCATATGAGAGACGCTGTCCGGTACAAAGTTC	854
Qy	2009	TATCTTTCCTTTGAAATTAATTCACACAAAGATTAACATCAGGAAAGCTAT---ACAAT	2065
Db	855	TATCTGGCCTTCGAGAACTCTCTTGACACCCGACTTACATATCCGAGAAAGCTGTGGAGAAC	914
Qy	2066	GCTTTTCTGGCTGGCTGTGTACTGTGTCTGTGTCTGTGGGACCATTAGGGAAAAATAGAGAA	2125
Db	915	GCCCTGGAGGCTGTGGCCGTGGCCGTGGTGTGGCCGACGAGAAAGCAACTACGAGAG	974
Qy	2126	TATATTCAGAGATTTATTCATTCATATGTGGAGATTTATTAATCTCCAGTGAAGTACGA	2185
Db	975	TTCTCGCACACCCGACGGCTTCATCCACGTGAGACATTCCTCAAGGCCCAAGGACCTGGCC	1034
Qy	2186	AAGTATCTGAAGAGAGACGACCAAAACAATAACTTATACCTTAGTTACTTTAATGAGAG	2245
Db	1035	CGGTAAGTGCAGAGAGCTGGACAAAGACACGCCGCTACCTGAGTACTTTCGCTGGCGG	1094
Qy	2246	AAG 2248	
Db	1095	GAG 1097	
RESULT 13			
AT61680	AAT61680 standard; DNA; 1654 BP.		
XX	AA61680;		
XX	AC		
XX	AT	19-JUN-1997 (first entry)	
XX	DE	Human alpha(1,3)-fucosyltransferase (Fuc-TV1) DNA.	
XX	KW	Alpha(1,3/1,4)-fucosyltransferase; Lewis enzyme; Fuc-TV1;	
XX	KW	glycosylation; oligosaccharide; blood group; ss.	
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
XX	FT	129..1208	
XX	FT	/*lag- a	
XX	XX	MO9709421-A1.	
XX	PD	13-MAR-1997.	
XX	PF	06-SEP-1996; 96WO-US13816.	
XX	PR	08-SEP-1995; 95US-0525058.	
XX	PA	(UNMT) UNIV MICHIGAN.	
XX	PI	Legault DJ, Lowe JB;	
XX	DR	WPI: 1997-192897/17.	
XX	DR	P-PSDB; AAM13643.	
XX	XX	New recombinant fucosyltransferase proteins - useful for modifying	
XX	XX	cell surface oligosaccharide structures	

Db 571 GACCTGCTGGCTGGCGGTGTCTCAACTGGAAGCCGAGCTACAGCCAGGGTGGCTCTAC 630
QY 1907 AATGACCTAAGCAAAAGCATTTGAATTCATCTACGGGCGACATTTGGAGATATGTC 1966
Db 631 CAGAGCTGAGGCTCATCTCAAGGTGAGCTGACGAGCGCTCCACAAGCCCGTGGCC 690
QY 1967 AATGATAAAAATTTGATTTCTTCCATATCTGCTTGAATTTTATCTTCTTGAAT 2026
Db 691 AAGGGAGACA--TGATGAGAGAGCGCTCCGCTACAGTTCTACCTGGCTTCGAGAAC 747
QY 2027 TCAATCCACAAGATTAACATACGGAAGCTAT---ACAATGCTTTTCTGGCTGCTC 2083
Db 748 TCTTTCACCCCGACTACATACCGGAAGCTGTGAGAGAACCCCTGGAGGCTGGGCC 807
QY 2084 GTACCTGTGTTCTGTGGACATCTAGGGAATAATATATATTCAGCAGATTCA 2143
Db 808 GTGCGCGTGTGCTGGGCCCCAGACAGCACTACGAGAGGTCTCTGCAACCCGACGCC 867
QY 2144 TTCAATTCATGTGGAATTTAATTAATCTCCAGTACGACCAAGATATCTGAAGAGTC 2203
Db 868 TTCAATTCATGTGGAATTTAATTAATCTCCAGTACGACCAAGATATCTGAAGAGTC 927
QY 2204 GACAAAACATTAAGTTATACCTTAAGTTTAACTGGAGAG 2248
Db 928 GACAAAGACACGCCCGCTACCTGAGCTACTTTCGCTGGCGGAG 972

RESULT 15

AAT80111
ID AAT80111 standard; cDNA; 2042 BP.

XX AAT80111;

DT 01-DEC-1997 (first entry)

XX Fucosyltransferase III coding sequence.

XX P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;
XX calcium-dependent carbohydrate binding protein; platelet; endothelium;
XX thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;
XX inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
XX disseminated intravascular coagulation; rheumatoid arthritis; antibody;
XX adult respiratory distress syndrome; tumor metastasis;
XX fucosyltransferase III; acetylglucosaminyltransferase;
XX autoimmune disorder; inflammatory disorder; ss.

OS Homo sapiens.

PN W09706176-A2.

XX 20-FEB-1997.

XX 02-AUG-1996; 96WO-US12820.

XX 17-MAY-1996; 96US-064802.

XX 03-AUG-1995; 95US-0510920.

XX 15-MAY-1996; 96US-0017794.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Cummings RD, McEever RP, Moore KL;

XX WPI; 1997-154206/14.

XX New O-glycan(s) derived from P-selectin glyco:protein ligand
XX (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
XX for treating inflammatory or autoimmune disorders or tumors

XX Disclosure; Page 80-81; 96pp; English.

XX This sequence represents the coding sequence for the human
XX fucosyltransferase III. The encoded protein, and the core 2 beta-6-N-
XX acetylglucosaminyltransferase encoded by AAT80112 are needed for the

CC expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in
CC mammalian cells. P-selectin is a calcium-dependent carbohydrate binding
CC protein expressed on the surfaces of activated platelets and endothelium
CC in response to thrombin and other agonists. PSGL-1 (see AAMW26174) is a
CC high affinity P-selectin ligand produced by leukocytes. Binding of
CC P-selectin to PSGL-1 is calcium ion dependent and is abolished by
CC treatment of the ligand with sialidase. PSGL is a homodimer, with a
CC highly extended extracellular domain, which is a feature of mucin-like
CC proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of
CC the invention (see AAMW26175-W26188). PSGL-1 is also targeted by the
CC O-glycans, and O-glycanated products of the invention. The products are
CC used for inhibiting the binding of P-selectin and other selectins to
CC PSGL-1. They can be used for the treatment of inflammation, injury
CC resulting from ischaemia and reperfusion, bacterial sepsis and
CC disseminated intravascular coagulation, adult respiratory distress
CC syndrome, tumor metastasis, rheumatoid arthritis, atherosclerosis and
CC other autoimmune or inflammatory disorders. The products can also be used
CC in the production of antibodies and in diagnostic applications.

SQ Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

Query Match 4.0%; Score 112.2; DB 18; Length 2042;

Best Local Similarity 51.6%; Pred. No 7.9e-13;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 1559 TGGCATCTCAAGACGAGACGCTTCACTGTACACAAATCCCATCAGTTCTGATCTAC 1618
Db 343 TGGCATCTCAAGACGAGACGCTTCACTGTACACAAATCCCATCAGTTCTGATCTAC 402
QY 1619 CGAGACATCAGTTGGG---TCTGACAAATTTACCTGACAGCTAGGCCATCTTCAG 1675
Db 403 TGGGATATCATGTCCAAACCTTAAGTCAAGCCTTCCCTCCCGAGCGGCGAGGGGAG 462
QY 1676 AATGATTTGATGATGATTTGGAATGGAATCAACCACTCACACCTCCCAAAAGAGTGCAT 1735
Db 463 CGCTGATCTGTTCAACTGTGGAGCCACCCCTTACCTGCGAGACCTGGAAGCTGTGAG 522
QY 1736 CACTTGTTAACCTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACT 1795
Db 523 AGATCTTCAATCTCAACATGCTCTTACCGGAGGACTCCGACATCTTCAACGCCCTTAC 582
QY 1796 TTTTGAAGGTTAAGCAACAAATCCCTTCG-----GTTGAAGTCCCAAGCAAGAG 1846
Db 583 TGGCTGAGGCGGTGTGTCGGCCAGCTGCCACCAACCGCTCAACCTCTCGCCAGAGCC 642
QY 1847 AATGTTGTGTCGGGTGTGAGTAAGTGAACCTGAGACCTGAGATCCAGAGCAAGTATTAC 1906
Db 643 GAGCTGTGCTGTGGCGGTGTGTCMACTGGAACCGGACTCAGCCAGGCTGACTAC 702
QY 1907 AATGACCTAAGCAAAAGCATTTGAATTCATCTACGGGCGACATTTGGAGATATGTC 1966
Db 703 CAGAGCTGAGGCTCATCTCAAGGTGAGCTGACGAGCGCTCCACAAGCCCGTGGCC 762
QY 1967 AATGATAAAAATTTGATTTCTTCCATATCTGCTTGAATTTTATCTTCTTGAAT 2026
Db 763 AAGGGAGACA--TGATGAGAGAGCTGTCCCGGTACAAAGTTCTACCTGGCTTCGAGAAC 819
QY 2027 TCAATCCACAAGATTAACATACGGAAGCTAT---ACAATGCTTTTCTGGCTGCTC 2083
Db 820 TCTTTCACCCCGACTACATACCGAGAGAGCTGTGAGAGAACCCCTGGAGGCTGGGCC 879
QY 2084 GTACCTGTGTTCTGTGGACATCTAGGAAATATATATATATATATATATATATATAT 2143
Db 880 GTGCGCGTGTGCTGGGCCCCAGACAGCACTACGAGAGGTCTCTGCAACCCGACGCC 939
QY 2144 TTCAATTCATGTGGAATTTAATTAATCTCCAGTACGACCAAGATATCTGAAGAGTC 2203
Db 940 TTCAATTCATGTGGAATTTAATTAATCTCCAGTACGACCAAGATATCTGAAGAGTC 999
QY 2204 GACAAAACATTAAGTTATACCTTAAGTTTAACTGGAGAGAG 2248
Db 1000 GACAAAGACACGCCCGCTACCTGAGCTACTTTCGCTGGCGGAG 1044

Tue Oct 8 10:18:24 2002

us-09-744-748-4.rng

Page 16

Search completed: October 5, 2002, 23:12:59
Job time : 476.687 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 2362.78 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-4

Perfect score: 2822 1 aagcttgctactcagagag.....aacatgttcgcatgaattc 2822

Sequence: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Scoring table: 13736207 seqs, 6748477542 residues

Searched: Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_estbhum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603.2	21.4	716	9	AV725071 AV725071
2	575.6	20.4	697	9	BB397507 BB397507
3	573.2	20.3	579	10	BM021081 BM021081
4	557.8	19.8	561	12	AQ321598 RPT111-93
5	492.8	17.5	607	10	BI102760 BI102760
6	448	15.9	501	12	B41073 B41073
7	409.2	14.5	649	9	AV338075 AV338075
8	296.8	10.5	548	12	AQ772584 AQ772584
9	292.6	10.4	1101	12	CNS059ST CNS059ST
10	280.4	9.9	648	9	BB595164 BB595164
11	269.6	9.6	625	9	BB649967 BB649967
12	251	8.9	767	9	BB640911 BB640911
13	227.6	8.1	568	10	BM021317 BM021317
14	184.2	6.5	628	9	BB651409 BB651409
15	177	6.3	570	10	BM082484 BM082484
16	170.2	6.0	1100	12	CNS0579Z CNS0579Z
17			1016	12	CNS0429P CNS0429P

18	165	5.8	844	12	CNS018V8	AL168808 Tetraodon
19	157.4	5.6	887	12	CNS0200S	AL208693 Tetraodon
20	155.4	5.5	651	10	BU036855	BU036855
21	151.6	5.4	666	10	BU050300	BU050300
22	151.2	5.4	711	12	CNS034XF	AL228012 Tetraodon
23	146.6	5.2	640	10	BU073227	BU073227
24	146	5.2	533	10	BU040066	BU040066
25	144.6	5.1	670	9	BB626126	BB626126
26	138	4.9	199	9	BB398103	BB398103
27	125.2	4.4	674	10	BB605503	BB605503
28	120.6	4.3	971	12	CNS047IV	AL278032 Tetraodon
29	120.4	4.3	615	10	BI226811	BI226811
30	119.8	4.2	284	9	BB270301	BB270301
31	117.8	4.2	1023	12	CNS02ANG	AL188773 Tetraodon
32	115.6	4.1	317	9	BB272545	BB272545
33	106.2	3.8	347	10	M78549	M78549 EST00697 Fe
34	103	3.6	702	10	BU043578	BU043578
35	102.8	3.6	487	10	BB63396	BB63396
36	96.2	3.4	1101	12	CNS0039G	AL210840 Tetraodon
37	93.8	3.3	863	12	CNS02ROF	AL063921 Drosophila
38	90.4	3.2	720	10	BF180380	BF180380
39	90	3.2	551	10	BU035701	BU035701
40	89.4	3.2	650	9	AV375511	AV375511
41	89	3.2	1101	12	CNS00EVL	AL069706 Drosophila
42	87.4	3.1	601	10	BU036679	BU036679
43	87.4	3.1	605	10	BU028804	BU028804
44	87	3.1	1056	12	CNS01VL7	AL169252 Tetraodon
45	86.8	3.1	304	9	AU081478	AU081478

ALIGNMENTS

RESULT 1
AV725071
LOCUS AV725071 716 bp mRNA linear EST 16-OCT-2000
DEFINITION AV725071 HTC Homo sapiens cDNA clone HTCLB02 5', mRNA sequence.
ACCESSION AV725071
VERSION AV725071.1 GI:10830091
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1 (bases 1 to 716)
Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,U., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
Homo sapiens cDNA HTC clones

JOURNAL
Unpublished (2000)
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..716
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTCLB02"
/clone_lib="HTC"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 219 a 113 c 133 g 249 t 2 others
ORIGIN

	Query Match similarity	21.4%	Score	603.2	DB 9:	Length	716;			
	Best Local Similarity	99.2%	Pred.	0.2	8e-78;					
	Matches	605;	Conservative	0;	Mismatches	5;	Indels	0;	Gaps	0;
OY	1947	AAGCATTTGGGAAATATGTCAATGATAAATAATTGATCTGTACATATCTGCTGTAAAT	2006							
Db	1	AAGCATTGGGAATATGTCAATGATAAATAATTGATCTGTACATATCTGCTGTAAAT	60							
OY	2007	TTTATCTTCTCTTGAAAATTCATATCCACAAGATTACATCACGAAAAAGCTATACAAATG	2066							
Db	61	TTTTATCTTCTTCTTGAAAATTCATATCCACAAGATTACATCACGAAAAAGCTATACAAATG	120							
OY	2067	CTTTTCTCGCTGCCTGCTGTACTGTGTGTTCTGGGACCACATCTGAGGAAAACTATAGAAAT	2126							
Db	121	CTTTTCTCGCTGCCTGCTGTACTGTGTGTTCTGGGACCACATCTGAGGAAAACTATAGAAAT	180							
OY	2127	ATATTCACGACGATATCTTCATTCATCAATGTAAGTAATTAATCTGCCAGTAGTAGTACCAA	2186							
Db	181	ATATTCACGACGATATCTTCATTCATTCATGTAAGTAATTAATCTGCCAGTAGTAGTACCAA	240							
OY	2187	AGTATCTGAAGAAGTGCACAAAAACAATAAGTATATACCTTAGTACTTTAACTGAGGA	2246							
Db	241	AGTATCTGAAGAAGTGCACAAAAACAATAAGTATATACCTTAGTACTTTAACTGAGGA	300							
OY	2247	AGCATTTTCACGTAAATCTTCCACAGATTTTGGGAATCACATGCATGTTGGCTTGCCATC	2306							
Db	301	AGGATTTTCACGTAAATCTTCCACAGATTTTGGGAATCACATGCATGTTGGCTTGCCATC	360							
OY	2307	ATGTAAAGGACATCAAGAAATTAAGTCTGTGTGGTAAATTAGAGAAATGGTTTGGAAAT	2366							
Db	361	ATGTAAAGGACATCAAGAAATTAAGTCTGTGTGGTAAATTAGAGAAATGGTTTGGAAAT	420							
OY	2367	AAAAATTTTTCATCCTTGACACCTGGATTAATATTTTGGATGAAATATCATCATCAGATNTG	2426							
Db	421	AAAAATTTTTCATCCTTGACACCTGGATTAATATTTTGGATGAAATATCATCATCAGATNTG	480							
OY	2427	AGGATTAAGAAGATGACATATCTACTTTTGTGTCCACATTTATTTTATACACCCTCTC	2486							
Db	481	AGGATTAAGAAGATGACATATCTACTTTTGTGTCCACATTTATTTTATACACCCTCTC	540							
OY	2487	TAGGGTAACTGTATATATTTTGGTGAGATTTTAAAGCTCAGCATGAGCAATCATTCOA	2546							
Db	541	TAGGGTAACTGTATATATTTTGGTGAGATTTTAAAGCTCAGCATGAGCAATCATTCOA	600							
OY	2547	TTTCGGTTTAA 2556								
Db	601	TTNCGTTTAA 610								
RESULT 2										
BB397507										
LOCUS										
DEFINITION	BB397507 RIKEN full-length enriched, ES cells Mus musculus cDNA clone Cj330004Cl9.3 similar to AB015426 Mus musculus Pcdy mRNA for alpha1,-3-fucosyltransferase IX, mRNA sequence.	697 bp	mRNA	linear	EST 24-OCT-2001					
ACCESSION	BB397507									
VERSION	BB397507.2									
KEYWORDS	B397507.2 GI:16411808									
SOURCE	EST.									
ORGANISM	house mouse.									
	Mus musculus									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
REFERENCE	1 (bases 1 to 697)									
AUTHORS	Arkawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.A., Kouda,									
	Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Koude,									
	'M., Koye,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y.,									
	Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sekai,K., Sano,H., Sasaki,H.,									
	D., Shibata,K., Shimaguchi,A., Shiraki,I., Sogabe,Y., Suzuki,H., Tagami,M.,									
	Tagawa,M., Tagawa,A., Takahashi,F., Takeida,I., Tanaka,T., Toyota,T.,									
	Muramatsu,M. and Hayashizaki,Y.									
	RIKEN Mouse ESTs (Arkawa,T., et al. 2001)									
TITLE	Unpublished (2001)									
JOURNAL	On July15, 2000 this sequence version replaced gi:9216903.									
COMMENT										

Contract: Yoshida Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res., 10 (10), 1617-1630 (2000)
Wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E.,
Matsubara, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
, S., Kawai, T., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res., 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamane, A., Aizawa
, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome, 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Location/Qualifiers
1. 697
/organism="Mus musculus"
/strain="CS7Bl/6J"
/db_xref="taxon:10090"
/clone="C330004C19"
/clone_idb="RIKEN full-length enriched, ES cells"
/cell_type="ES cells"
/lab_host="SOLR"
/note="Site.1: XhoI; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAAGAGGAGGATGATCAAGAGGCTTTTGTTCCTTTGAGGCTAAGCA 1811
3'] cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 5.0 and subtraction to Rot = 25.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GAGAGGAGATTCGATTCAGATTCCAAGGCTTTTGTTCCTTTGAGGCTAAGCA 1811
3']


```

|||||
Db 39 TATTCCTTCATATCATTTGTGAACATTTGCTTTTTCACA 2
|||||
RESULT 4
LOCUS AO321598/c 561 bp DNA linear GSS 06-MAY-1999
DEFINITION RPl11-93E9.TV RPl1-11 Homo sapiens genomic clone RPl1-11-93E9, DNA
sequence.
ACCESSION AO321598
VERSION AO321598.1 GI:4054332
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 561)
Adams M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Sub, E., Wible, C., de Jong, P., and Ventner, J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
Other_GSSs: RPl11-93E9.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are derived from the human BAC library RPl1-11. For BAC
library availability, please contact Pieter de Jong
(pjeter@edlong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (Info@resgen.com). BAC end search page:
http://www.ligr.org/tdb/humgen/bac_end_search.html
Seq primer: 17
Class: BAC ends.
FEATURES
source location/Qualifiers
1..561
/organism="Homo sapiens"
/db_xref="GDB:7535432"
/db_xref="taxon:9606"
/clone="RPl1-11-93E9"
/clone_11b="RPl1-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPl11 Human Male BAC Library"
BASE COUNT 200 a 92 c 89 g 180 t
ORIGIN
Query Match 19.8%; Score 557.8; DB 12; Length 561;
Best Local Similarity 99.6%; Pred. No. 1.1e-71;
Matches 559; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 717 TGCATTTTTCACAGATGATTAATGATTCGATTCGCAACTGTGAAGTAATTTAT 776
|||||
Db 561 TGCATTTTTCACAGATGATTAATGATTCGATTCGCAACTGTGAAGTAATTTAT 592
|||||
Db 777 AAGGAGAAAAAACAATACGTTTAAACAGTTTCTTAAAGCAATTCAT 836
|||||
Db 501 AAGGAGAAAAAACAATACGTTTAAACAGTTTCTTAAAGCAATTCAT 442
|||||
Db 837 CTGACCCACAAGTTAGTACAGTCAATATGAGTAATTAAGTATATCAACT 896
|||||
Db 441 CTGACCCACAAGTTAGTACAGTCAATATGAGTAATTAAGTATATCAACT 382
|||||
Db 897 GGGAGATGTTCAAAATCGATGTTGTTACAAAATCTTAAATTTAGTCATGCTTTA 956
|||||
Db 381 GGGAGATGTTCAAAATCGATGTTGTTACAAAATCTTAAATTTAGTCATGCTTTA 322
|||||
Db 957 GCGATGACAGTCTAGGTTGCGCACTTTCTACAAACAATAAAGTGAACCTGTAAAG 1016
|||||

```

```

Db 321 GCGATGACAGTCTAGGTTGCGCACTTTCTACAGCAATTAAGTACAACTTAAG 262
|||||
Db 1017 ATAGCATTTGAAAAAAGCTTACTGTTTGAAGATATGCTGACTTTTACAGAGTTT 1076
|||||
Db 261 ATAGCATTTGAAAAAAGCTTACTGTTTGAAGATATGCTGACTTTTACAGAGTTT 202
|||||
Db 1077 TGGAAACCTCTTTTGAAGAGCTGTTCTGCTTGGTGGAGAAAGTTCCTCTCAAT 1136
|||||
Db 201 TGGAAACCTCTTTTGAAGAGCTGTTCTGCTTGGTGGAGAAAGTTCCTCTCAAT 142
|||||
Db 1137 GATTGATTAACACTGAAAGAAAGAAATGCTGTGATTAATGATTAATGATTAAT 1196
|||||
Db 141 GATTGATTAACACTGAAAGAAAGAAATGCTGTGATTAATGATTAATGATTAAT 82
|||||
Db 1197 GATTGATTAATATCTCCAAATATATTTGATTTCTGCTTCAATCCACCGCTACCTCC 1256
|||||
Db 81 GATTGATTAATATCTCCAAATATATTTGATTTCTGCTTCAATCCACCGCTACCTCC 22
|||||
Db 1257 CTCTGCTTCTTCTCTATTTTCG 1277
|||||
Db 21 CTCTGCTTCTTCTCTATTTTCG 1
|||||
RESULT 5
LOCUS B1102760 607 bp mRNA linear EST 26-JUN-2001
DEFINITION 602888338F1 NCI CGAP_K1d14 Mus musculus cdna clone IMAGE:5043609
5', mRNA sequence.
ACCESSION B1102760
VERSION B1102760.1 GI:14553653
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 607)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1119 row: 1 column: 10
High quality sequence stop: 606.
FEATURES
source location/Qualifiers
1..607
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5043609"
/clone_11b="NCI CGAP_K1d14"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: Kidney; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP library."
BASE COUNT 179 a 149 c 121 g 157 t
ORIGIN
Query Match 17.5%; Score 492.8; DB 10; Length 607;
Best Local Similarity 90.1%; Pred. No. 2.9e-62;
Matches 549; Conservative 0; Mismatches 58; Indels 2; Gaps 2;
Db 1511 CAGACCTTTGACCTTCAATCTCCAGCAAGATGTTCAACATCCAGATGCAATCTACA 1570
|||||
Db 1 CAGACCTTTGACCTTCAATCTCCAGCAAGATGTTCAACATCCAGATGCAATCTACA 60
|||||

```

OY	1571	ACGGAGCCGTTCACTGATGACAACAATAATCCCATGCAGTTCATCATCATTACCAGACATCACT	1630
Db	61	ACAGAACCGGTGA -TGATACAAACAATGCCATGGGCTCATTTACCATGAGACATCACG	119
OY	1631	TGGGATCTTACAAAATTTACTCCAGCAAGCTAGAGCCACCCTTCACAGAAATGATTTGGANG	1690
Db	120	TGGGATCTTACACTTACTTACTTACCTACGACAGGCCAAGCCACCTTTCAGAAAAATGA-TTGGANG	178
OY	1691	AATTTGGAATCACCAACACTCACATCCCCCCAAAAGAGTAGTGCAATTGAGCATGTTATACCTG	1750
Db	179	AATTTAGAGTACACCACTACACACCCCACAAAAGAGTGGCATTTGAACACTTGTTCAACCTG	238
OY	1751	ACTCTGACTTACCCGCCGTGATTCACAATATCCAAGTGCCTTATGCTTCTTGACGGTAGC	1810
Db	239	ACTCTTAATTTATCGCCGTGATTCACAATATCCAAGTGCCTTATGCTTCTTGACGGTAGC	298
OY	1811	ACAAATCCCTTGSTGTTTGAAGTGGCAAGCAAGAATAATGTGGTGGCTGGGTGGAGT	1870
Db	299	ACAAATCCCTTGTTGTITTTGAAGTGGCCACCAAGAGAAAGTGGTGGCTGGGTGGAGT	358
OY	1871	AACTGGAAACCTGAGCATGCGCAGAGTCAAGTATTAACAATGAGCTAAGCAAAAGACTGAA	1930
Db	359	AACTGGAAACCTGAGCATGCGCAGAGTCAAGTATTAACAAGACTCAGCAAGAGTATGAA	418
OY	1931	ATCCATACCTACGGGCAAGCATTTGGAGAAATATGTCATATGATAAAAATTTGATTCCTACC	1990
Db	419	ATCCACACTTATGAGGCCAAGCATTCGGAGAAATCGTGAAAGATAAAAATTCGATTTCCACC	478
OY	1991	ATATCTGCTGTGAATTTTATCTTCTCTTGAANAATCAATCACAAGATTAACATCAGC	2050
Db	479	ATATCTACTGTGAATTTTATCTTCTTCTTGAATTCATTCATCACAAATTAACATCACA	538
OY	2051	GAAAGCTATACAAATGCTTTTCTGGCTGAGCTGTACTCTGTTCTGTGGAGCATCTAGG	2110
Db	539	GAAAGCTCTACAAATGCAATTTTTGGCTGTTCAAGTACCTGTTCTGTGGAGCATCTAGG	598
OY	2111	GAAACATAT 2119 	
Db	599	GAAACATAT 607 	
RESULT 6	B41073/c	501 bp DNA linear GSS 18-OCT-1997	
LOCUS	DEFINITION	HS-1052-B2-G01-MF.abl CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate-Cr 774 Col=2 Row=N, DNA sequence.	
ACCESSION	B41073		
VERSION	B41073.1	GI:2545325	
KEYWORDS	GSS.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 501) Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S., Trafletoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E. Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Mahairas GG, Zackrone KD, Hood L University of Washington Seattle, WA 98195, USA Tel: (206) 616-8744 Fax: (206) 685-7301 Email: kzackrone@u.washington.edu Sequence Tagged Connector Plate: Cr 774 row: N column: 2 Class: BAC ends High quality sequence stop: 501. Location/Qualifiers 1..501 /organism="Homo sapiens"		
FEATURES	source		

[illegible]

(pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 1144 row: G column: 18
Seq primer: T7
Class: BAC ends
High quality sequence stop: 548.
Location/Qualifiers

FEATURES

source

1..548

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate-1144 Col-18 Row-G"

/clone_1lb="Rpci-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

BASE COUNT 190 a 94 c 89 g 168 t 7 others
ORIGIN

Query Match

Best Local Similarity 14.5%; Score 409.2; DB 12; Length 548;

Matches 426; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 2366 TAAATTTTTCATCATCTGCACACTGATTAATATTTTGTGATGATATCCACAGATT 2425
|||||
Db 547 TAAATTTTTCATCATCTGCACACTGATTAATATTTTGTGATGATATCCACAGATT 488
|||||

QY 2426 GAGGATGAAGAGATGCAACATAGTCTTGTGCACAAATTTATTTATCACCCTC 2485
|||||
Db 487 GAGGATGAAGAGATGCAACATAGTCTTGTGCACAAATTTATTTATCACCCTC 428
|||||

QY 2486 CTGAGGTAACGCTATATTTTGTGAGAGATTTTAAAGCTGACATGACATCATTC 2545
|||||
Db 427 CTGAGGTAACGCTATATTTTGTGAGAGATTTTAAAGCTGACATGACATCATTC 368
|||||

QY 2546 ATTCGGTTTAAATTTCCGTATATACCTAATATATGTCACATGAGAGATTTATTC 2605
|||||
Db 367 ATTCGGTTTAAATTTCCGTATATACCTAATATATGTCACATGAGAGATTTATTC 308
|||||

QY 2606 TCATATCATTTTAAACATGCTTTTTCACATTTTGTAGTGTCCATATATGTAAGCT 2665
|||||
Db 307 TCATATCATTTTAAACATGCTTTTTCACATTTTGTAGTGTCCATATATGTAAGCT 248
|||||

QY 2666 GGTGTTGATATTTTTCACATGCTGTTAATCTTTTGGGAAAGAGATG 2725
|||||
Db 247 GGTGTTGATATTTTTCACATGCTGTTAATCTTTTGGGAAAGAGATG 188
|||||

QY 2726 CACATCTTAAAGTAAAAATTTTCACTAATGATTAACAATGCTAGTTCCAACTTTGCA 2785
|||||
Db 187 CACATCTTAAAGTAAAAATTTTCACTAATGATTAACAATGCTAGTTCCAACTTTGCA 128
|||||

QY 2786 TACTATTAACAGAGAAACATGCTGATTAATTC 2822
|||||
Db 127 TACTATTAACAGAGAAACATGCTGATTAATTC 91
|||||

RESULT 9

CNS059ST/c

LOCUS

DEFINITION

1101 bp

DNA

linear

GSS

26-MAY-2000

sequence.

Tetradon nigroviridis genome survey sequence T3 end of clone

045922 of library A from Tetradon nigroviridis, genomic survey

ACCESSION

AL327638

VERSION

GI:8221227

KEYWORDS

GSS; genome survey sequence.

SOURCE

Tetradon nigroviridis.

ORGANISM

Tetradon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetradontiformes; Tetradontidae; Tetradon.
1 (bases 1 to 1101)
Roest-Crolius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the
freshwater purifier Tetradon nigroviridis

JOURNAL Unpublished
2 (bases 1 to 1101)
Roest-Crolius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C.,
Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.

AUTHORS Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence

JOURNAL Unpublished
3 (bases 1 to 1101)
Roest-Crolius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C.,
Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.

REFERENCE Genoscope.
Direct Submission
Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetradon>.

COMMENT JOURNAL
TITLE Direct Submission
Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetradon>.

FEATURES
source
Location/Qualifiers
1..1101
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone="045P22"
/clone_1lb="A"

BASE COUNT 287 a 249 c 269 g 278 t 18 others
ORIGIN

Query Match 10.5%; Score 296.8; DB 12; Length 1101;
Best Local Similarity 58.3%; Pred. No. 5.8e-34;

Matches 603; Conservative 0; Mismatches 423; Indels 9; Gaps 5;

QY 1338 TTATCGTGGGCTGTTTCAAGGATGCTCTCTCTTATACATCAACCTTACACAGCTGGA 1397
|||||
Db 1064 TTATCGTGGGCTGTTTCAAGGATGCTCTCTCTTATACATCAACCTTACACAGCTGGA 1006
|||||

QY 1398 TCTTCAGTCCATGGAATGACCGAGCTGTGTGAAATGAAATCTTTTCCACCA 1457
|||||
Db 1005 TCTTCAGTCCATGGAATGACCGAGCTGTGTGAAATGAAATCTTTTCCACCA 946
|||||

QY 1458 AAACGATATTTTATGAAATGATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1517
|||||
Db 945 AAACGATATTTTATGAAATGATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 889
|||||

QY 1518 TTGACCTTACATCTGCGCAAGCAATGTTCAACATCCAGAGATGCCATCCACAGGACC 1577
|||||
Db 888 TTGACCTTACATCTGCGCAAGCAATGTTCAACATCCAGAGATGCCATCCACAGGACC 829
|||||

QY 1578 GTTCACTGTAACAACAAATCCATGCAAGTTCTGATCCATACCGAGACATCAAGTTGGATC 1637
|||||
Db 828 GTTCACTGTAACAACAAATCCATGCAAGTTCTGATCCATACCGAGACATCAAGTTGGATC 769
|||||

QY 1638 TGACAAATTTTACCTGCAAGCAAGTACCGCCACCTTCCAGAAATGATTTGGATGAATTGG 1697
|||||
Db 768 TGACAAATTTTACCTGCAAGCAAGTACCGCCACCTTCCAGAAATGATTTGGATGAATTGG 709
|||||

QY 1698 AATCACAACATGCAACTGCCCAAAAGAGGCAATGAGCAATGTTTACCTGACATCTGA 1757
|||||
Db 708 AATCACAACATGCAACTGCCCAAAAGAGGCAATGAGCAATGTTTACCTGACATCTGA 649
|||||

QY 1758 GTTACCGCCGATTCAGATATCCAAAGTGCCTTATGCTCTTGAACGTTAAGCAAAATC 1817
|||||
Db 648 GTTACCGCCGATTCAGATATCCAAAGTGCCTTATGCTCTTGAACGTTAAGCAAAATC 589
|||||

QY 1818 CCTGCTG---TTTGAAGGCCAAGCAAGAAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1874
|||||
Db 588 CCTGCTG---TTTGAAGGCCAAGCAAGAAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 529
|||||

OY	1875	GGAACCCAGCATTGCGCAGATCAACTATTTCAATGGACCTAAACAAAAGCATGTGAATTC	1934
Db	528	GGAACCGAAGAATGCACGATGTAATTCTCAACAGATTTGTACMAAGCATGTTGAGTCC	469
OY	1935	ATTACCGAGGGCAGCATTTTGSGAATAATGTCAATGATTAATAATTGATTCCTACCAT	1994
Db	468	AACCGATATGACGAGCATTCGGSCAACACTTCCAGATGAAGAGACCTTACCACACATCG	409
OY	1995	TGCTCTTGAAATTTATCTTCTCTTGAAAAATGCAACCCACAAGATTTCAATCGAGAAA	2054
Db	408	CCAGCTGCAGATGTACTCCTGTCAITTTGGAAATCATGCCAACAGACTACTTTACTGAAA	349
OY	2055	AGCTATATCAATGCTTTTTCTGGCTGGCTGTACCTGCTTTGCTTGGGACCATCTTAGGAAA	2114
Db	348	AACTGTACACCCCACCTGCTGTGGGACAGTCCAGAGGTCTTGGGGCCACCTCGGAGGA	289
OY	2115	ACGATGACAAATTATTTTCACAGCAGATTCATTCATCTGAGGAATTTATATCTCTCCA	2174
Db	288	ACTACGABAAATATCATCCACAGGAGATGACCTTTATCCAGGGAAAGCATTTGGCTCACCA	229
OY	2175	GTGAGCTAGCAAATGATCTGAMAGGAATGACGACAAAACAATAAGATATACCTTAGTACT	2234
Db	228	AGAGAGCTTGCTGACTACTATTTTACTGCTCTCGAGCAAAAATVAGGAAATGTACCTTTAAATTT	169
OY	2235	TTTAACTGGAGGAAGATTTTCACTGTAAATCTTCCACAGATTTTGGAGATCAATCATCATGTT	2294
Db	168	TTTAAATGTCGCAADACACTTTAAAGCAGAGGGCAGTGCATTTTCCCGGAGG-CACCTGTC	110
OY	2295	TGGCTTTCGCATCATGTGAAAAAGGCATCAAGAATTA-AGTCTGTTGGTATTTAGAGAAA	2353
Db	109	AGGCTTTGTGATTTACGTCCAGAGGCGNTINACAAATTAACAGATTCAGTGCACCTTANACAG	50
OY	2354	TGGTTTGGAAATTAA 2368	
Db	49	TGGTACTGGGGGTGA 35	

RESULT 10				
BBS95164				
LOCUS	BB955164	648 bp	mRNA	linear
DEFINITION	BR955164 RIKEN full-length cDNA clone C030004.N01 S ^t , mRNA sequence.			
ACCSSION VERSION	BB955164 BB955164			
KEYWORDS	musculus cdna clone C030004.N01 S ^t , mRNA sequence.			
SOURCE	EST.			
	house mouse.			

REFERENCE AUTHORS

TITLE
JOURNAL
COMMENT

Eukaryota; Vertebrata; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 648)
Arkawa, T., Carnicini, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koyota, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tezumi, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arkawa, T., et al. 2001)
Unpublished (2001)
On Nov 30, 2000 this sequence version replaced gi:11491766.
Contact: Yoshinobu Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9212
Fax: 81-45-503-9216
E-mail: genome.res@gsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Carnicini, P., Shibata, Y., Hayashi, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cDNA-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

Wagli, K., Fujikawa, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matsuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

Riken integrated sequence analysis (RISA) system-384 format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shibaayawa, A., Saito, K., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research group in Riken Genetic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

1. 648 Location/Qualifiers

FEATURES

source

BASE COUNT	141 a	206 c	138 g	163 t	ORIGIN
Query Match	10.4%	Score 292.6;	DB 9;	Length 648;	
Best Local Similarity	90.2%	Pred. No. 2,8e-33			
Matches 313;	Conservative	0;	Mismatches 34;	Indels	0;
0Y	1281	GAATAATATATACATCAATCAAGAAATCTTGGCCCATTTTAAATATGTGCGATT	1340		
Db	302	GAATAATATATACATCAACATCCAAAGCATTTCTGCCCATTTTAAATGCTGCGATCA	361		
0Y	1341	TGCTGGGCTGTTTATATGCGATGTCTTTCATTTCATCAACCTACCAACGCTGGATCT	1400		
Db	362	TGCTGGGCTGCTTATGCGATGTCTGCTCATTTTATCAACGCCCAACACGCTGGATCT	421		
0Y	1401	TGATGTCATATGGAATCAGCCAGCTGTGTGCGAAATGAAAAACCTCTTTTCCACAAAA	1460		

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	422	TCAGCCCAATGGAGCTGCAGAGTTCTGTGCTGAGAAAATGAAAAATTTCTTCTCCACAAAAA	481					
Qy	1461	CTGATATATTTTANGAAAACCTACTATTCGTCTGTGTGGGTGTGGCCATTGTGGCAGACCTTTG	1520					
Db	482	CTGATATATTTTANGAAAACCTACTATTCGTCTGTGTGGGTGTGGCCATTGTGGCAGACCTTTG	541					
Qy	1521	ACCTTACATCCCTGGCAGAACACTGTTCACATCCCAAGATGTCATCTCCACACGGACGCTT	1580					
Db	542	ACCTTACATCCCTGGCAGAACACTGTTCACATCCCAAGATGTCATCTCCACACAGACCGCT	601					
Qy	1581	CACCTGTACACAAATCCCATCTGACAGTTCATTCATCCATCCACGACAGATC	1627					
Db	602	CATGTGTACACAAATCCCATCTGACAGTTCATTCATCCATCCACATTCAGATC	648					
RESULT 11								
LOCUS	BB649967							
DEFINITION	BB649967	625 bp	mRNA	linear	EST 26-OCT-2001			
ACCESSION	BB649967							
VERSION	BB649967							
KEYWORDS	BB649967.1	GI:16484222						
SOURCE	house mouse.							
ORGANISM	Mus musculus							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,T., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.							
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)							
JOURNAL	Unpublished (2001)							
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagii,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res 10 (11), 1757-1771 (2000) Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamane,K., I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues. Location/Qualifiers 1. 625							

Query Match	Best Local Similarity	Matches	Score	DB	Length
Query Match	Best Local Similarity	Matches	Score	DB	Length
Matches 296; Conservative	91.9%; Pred. No. 1,7e-31;	0; Mismatches 26; Indels 0; Gaps 0;			
1281	GAATAATATATACATCAATCAATCAAAAGGAATTCCTGCCCATTTTAAATGTCGATTA	1340			
303	GAATAATATATACATCAATCAATCAAAAGGATTCCTGCCCATTTTAAATGTCGATTA	362			
1341	TCCTGGGCTGTTTTCATGATGATGTCCTTCATTTACATCAAACTACAAACAGTGGATCT	1400			
363	TCCTGGGCTGTTTTCATGATGATGTCCTTCATTTACATCAAACTACAAACAGTGGATCT	422			
1401	TCAGTCACATGGAATCAGCCAGCTCTGTGTCGAAATGAAATGAAATCTCTTTCCACAAA	1460			
423	TCAGTCACATGGAATGAGCTGTCGAAATGTCGTGTCGAAATGAAATGAAATCTCTTTCCACAAA	482			
1461	CTGATTTATTTAATGAACACTCTAATCTGCTGGGCTGTGGCCATTTGGGACAGACTTGG	1520			
483	CTGATTTATTTAATGAACACTCTAATCTGCTGGGCTGTGGCCATTTGGGACAGACTTGG	542			
1521	ACCTTACATCTCTGCCACAGCATGTTCAACATCCAAAGATGATCTCAACAGACGCTT	1580			
543	ACCTTACATCTCTCTGCCACAGCATGTTCAATATCCAAAGGATGATCTCAACAGACGCTT	602			
1581	CACGTACACAAATCCGATGC	1602			
603	CATGTACACAAATCCGATGC	624			
RESULT 12					
LOCUS	BB640911				
DEFINITION	musculus cDNA clone A730093G08 5', mRNA sequence.				
ACCESSION	BB640911				
VERSION	BB640911.1				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	1 (bases 1 to 767)				
AUTHORS	Arikawa, T., Carrino, P., Fukuda, S., Furuno, M., Hanaoka, T., Hara, A., Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Komori, H., Kouda				


```

OY 1521 ACCTACATCTCTGCGAAGCAATGTTCAATCCAAAG 1557
Db 579 ACCTACATCTCTGCGAAGCAATGTTCAATCCAAAG 615

RESULT 15
BM082484/c
LOCUS BM082484
DEFINITION 570 bp mRNA linear EST 14-NOV-2001
            t23c508.x1 Campbell zebrafish ovary Danio rerio cDNA clone 5306223
            3 similar to TR:088819 088819 ALPHAI.3-FUCOSYLTRANSFERASE IX.1;
            mRNA sequence.
ACCESSION BM082484
VERSION   BM082484.1 GI:16929414
KEYWORDS EST.
SOURCE    zebrafish.
ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            1 Cyprinidae; Danio.
            1 (bases 1 to 570)
            Clark, M., Johnson, S.L., Lebrach, H., Lee, R., Li, F., Marra, M., Eddy
            S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
            R., Steptoe, M., Theisling, B., Allen, M., Bowers, Y., Person, B.,
            Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rilter, E.,
            Korn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
            and Wilson, R.
            Washu Zebrafish EST Project 1998
            Unpublished (1998)
            Other ESTs: t23c508.y1
            Contact: Stephen L. Johnson
            Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbratishewatson.wustl.edu
            DNA Library constructed by library constructed by invitrogen and
            donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA).
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: the I.M.A.G.E. Consortium/LINL, send email to:
            linlimage@linl.gov
            High quality sequence stop: 419.
FEATURES
    source
        Location/Qualifiers
            1..570
                /organism="Danio rerio"
                /db_xref="taxon:7955"
                /clone="5306223"
                /clone_lib="Campbell zebrafish ovary"
                /sex="female"
                /dev_stage="4-5 months, 1 year and 2 years"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: ovary (pooled); Vector: PCMV-SPORT6; site: 1;
                Note: Site: 2; Sail; Whole ovaries collected from zebrafish
                aged 4-5 months, 1 year and 2 years. Oligo-dT primed,
                directionally cloned. Average insert size 2 kb. Library
                constructed by invitrogen and donated by R. Campbell
                (Marine Biology Laboratory, Woods Hole, MA)."
```

BASE COUNT 155 a 106 c 129 g 180 t

ORIGIN

```

Query Match 6.5% Score 184.2; DB 10; Length 570;
Best Local Similarity 59.2%; Pred. No. 1.5e-17;
Matches 337; Conservative 0; Mismatches 223; Indels 9; Gaps 1;

OY 1670 TTCGAGAAATGAGTTGGATGATTTGAAATCACCACACTACCTCCCAAAAGAGTGGC 1729
Db 570 TTCGAGAAATGAGTTGGATGATTTGAAATCACCACACTACCTCCCAAAAGAGTGGC 511
OY 1730 ATGAGCACTGTGTTAACTGACTGCTGATACCGCCGCTGATTCAGATATCCAAAGTGCTT 1789
Db 510 CTGGAAACCTGTTCACACACTCTGAGCTACACAGCAAGATGCTGATATTCAGTACGG 451
OY 1790 TATGGCTTCTTACGAGTACCAAAATCCCTTGTTGAAAGTCCCAAGCAAGAGAAA 1849
```

```

Db 450 ATGAGATTTAAAMACAGAGAAAGAACCCGCGATGATTTTGTCAATTCACAAAAGACAAA 391
OY 1850 TTGGTGTGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1900
Db 390 CTAGTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 331
OY 1901 TATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1960
Db 330 ATTTATGCGGAATTAAGCCGATACATCCATGATGATGATGATGATGATGATGATGATGATGAT 271
OY 1961 TATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2020
Db 270 TATCTAGACTATTAAGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 211
OY 2021 GAAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2080
Db 210 GAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 151
OY 2081 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2140
Db 150 ACTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 91
OY 2141 TATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2200
Db 90 GCCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 31
OY 2201 GTTCGCAAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2229
Db 30 CTAAATTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2
```

Search completed: October 6, 2002, 07:15:52
Job time : 2383.78 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 57.5794 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-4

Perfect score: 2822
Sequence: 1 aagcttctactacagagag.....aacatgttcgattgaattc 2822

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	127	4.5	1814	2	US-08-483-151-1
2	127	4.5	1814	5	PCT-US86-06427-1
3	113.4	4.0	1654	1	US-07-914-281-13
4	113.4	4.0	1654	1	US-08-393-246-13
5	113.4	4.0	1654	1	US-08-525-058A-13
6	113.4	4.0	1654	2	US-08-696-731-13
7	113.4	4.0	1654	4	US-09-042-531-13
8	112.2	4.0	1086	1	US-07-914-281-12
9	112.2	4.0	1086	1	US-08-393-246-12
10	112.2	4.0	1086	1	US-08-525-058A-12
11	112.2	4.0	1086	2	US-08-696-731-12
12	112.2	4.0	1086	2	US-09-042-531-12
13	112.2	4.0	2043	3	US-09-063-237-2
14	112.2	4.0	2043	1	US-07-914-281-1
15	112.2	4.0	2043	1	US-08-393-246-1
16	112.2	4.0	2043	1	US-08-525-058A-1
17	112.2	4.0	2043	1	US-08-696-731-1
18	112.2	4.0	2043	2	US-08-696-731-1
19	112.2	4.0	2043	4	US-09-042-531-1
20	112.2	4.0	2043	5	PCT-US81-00899-1
21	111.8	4.0	1316	1	US-07-914-281-10
22	111.8	4.0	1316	1	US-08-393-246-10
23	111.8	4.0	1316	1	US-08-525-058A-10
24	111.8	4.0	1316	2	US-08-696-731-10
25	111.8	4.0	1316	4	US-09-042-531-10
26	100.8	3.6	1068	5	PCT-US81-00899-13
27	100.8	3.6	1215	5	PCT-US91-00899-5

28	100.8	3.6	1488	1	US-07-914-281-9	Sequence 9, Appl
29	100.8	3.6	1488	1	US-08-393-246-9	Sequence 9, Appl
30	100.8	3.6	1488	1	US-08-525-058A-9	Sequence 9, Appl
31	100.8	3.6	1488	2	US-08-696-731-9	Sequence 9, Appl
32	100.8	3.6	1488	4	US-09-042-531-9	Sequence 9, Appl
33	100.8	3.6	2175	4	US-08-482-073-9	Sequence 9, Appl
34	100.8	3.6	2861	4	US-08-482-073-10	Sequence 10, Appl
35	100.8	3.6	3647	1	US-07-914-281-7	Sequence 7, Appl
36	100.8	3.6	3647	1	US-08-393-246-7	Sequence 7, Appl
37	100.8	3.6	3647	1	US-08-525-058A-7	Sequence 7, Appl
38	100.8	3.6	3647	2	US-08-696-731-7	Sequence 7, Appl
39	100.8	3.6	3647	4	US-09-042-531-7	Sequence 7, Appl
40	100.8	3.6	3647	5	PCT-US91-00899-4	Sequence 4, Appl
41	97.6	3.5	2134	5	US-08-483-151-3	Sequence 3, Appl
42	97.6	3.5	2134	5	PCT-US96-06427-3	Sequence 3, Appl
43	63	2.2	4673	5	US-07-638-431-1	Sequence 1, Appl
44	63	2.2	4673	5	PCT-US92-00018-1	Sequence 1, Appl
45	54.2	1.9	2042	2	US-08-933-821-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-08-483-151-1
; Sequence 1, Application US/08483151
; Patent No. 5858752
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Holgersson, Jan
; TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,151
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/278001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-483-151-1

Query Match 4.5%; Score 127; DB 2; Length 1814;
Best local Similarity 52.3%; Pred. No. 1.4e-18;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

0Y 1557 GATGCATCTCAACAGCGGCTTCCTGATCAACAAATCCATGATGATTCATTCATC 1616
DB 548 GCTGCCGTGTGATGCTAACCAGGCTCTGACCGCTGATGCTGTGCTGCTTCACACC 607
```

Tue Oct 8 10:18:24 2002

us-09-744-748-4.rml

Page 2

```
OY 1617 ACCGACATCATGTTGGATCTGACAAATTTACCTACGACAGCTAGGCCACCTCCACA 1676
      |||||
Db 608 ACCGAGCTGCACAAACCGGAAATCTCTCCATCCCTGGACGAGGAGCCACGACGACG 667
OY 1677 AATGATTTGATGAAATTTGAAATCACCACATCCACATCCCAAAAAGAGTGCATTGAGC 1736
      |||||
Db 668 CTTGGCTCTGGGCTCTCCATGGAATCGCCAGATATACCATGATGTCATCGCTCCGGG 727
OY 1737 ACTTGTAACTGACTGACTGACTGACGCGGTGATTCAGATATCCAAAGTCCATTAGGCT 1796
      |||||
Db 728 GCATCTTCAACTGGGTGGGTGAGTATGGGGTATTCAGATATCTTTGATCCCTACGGTC 787
OY 1797 TCTTGACGGTAGACACAAATCCCTTGTGTTGATGTCGACGCAAGCAAGAAATTTGCT 1856
      |||||
Db 788 GCTTGGAGCCCTCTCTCTGGGCC---CACATCCCACTACGCGCAAAAGCTGACGCGAGCTG 844
OY 1857 GCTGGCTTGTGATGATGAGAACCCCTGAGCATCCAGAGTCAAGATTTACATGAGCTAA 1916
      |||||
Db 845 CCGTGGGTGATGACAAATTTCCAGAGGCGAGCAGAGCTGACAAAGCTGACGCGAGCTGG 904
OY 1917 GCAAAAGCATTTGAATCCATCTACGAGGCAAGCATTTGGAGATATGTCAATGATATAA 1976
      |||||
Db 905 CCCCTCATCTGACAGGTGATGTGTCGTCGCGCCAGCGAGCGGCCCTATGCGCTAAT 964
OY 1977 ATTGATTCCTACATATCTGCTGTAATTTATGTTTCCTTTGAAAAATTCATTCACA 2036
      |||||
Db 965 GTCTGCTGCCACTTTGGCCCGGATACCGCTTCTACCTGGCTTTGAGAACTACAGCATC 1024
OY 2037 AGATTTACATCAGGAAAAAGCTAT---ACAATGCTTTTCTGGCTGGCTGTACCTGTG 2093
      |||||
Db 1025 GGGACTACATCAGTGAAGTGTGGCGCAATGCTCGGCGGCTGTCTGTACCCCTGG 1084
OY 2094 TTCTGGACCATCTAGGAAAACTAGTAATATATTCACAGATTCATTCAATTCATG 2153
      |||||
Db 1085 CGGTGGACCTCTCTGGGCGACCTACGAGAGCTTTGTGGCACCAATGCTTTGTACAG 1144
OY 2154 TGGAGATTTAATCTCTCCAGTAGAGCTAGC 2184
      |||||
Db 1145 TGGACGACTTCAGCTCTGCCGCTGAACTGCG 1175
```

RESULT 2

```
PCT-US96-06427-1
: Sequence 1, Application PC/TUS9606427
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation
: TITLE OF INVENTION: FOCOSITRANSFERASE GENES AND USES THEREOF
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/06427
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/483,151
: FILING DATE: 07-JUN-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Lech, Karen F.
: REGISTRATION NUMBER: 35,238
: REFERENCE/DOCKET NUMBER: 00786/278M01
: TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1814 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
PCT-US96-06427-1
```

```
Query Match 4.58; Score 127; DB 5; Length 1814;
Best Local Similarity 52.38; Pred. No. 1,4e-18;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
```

```
OY 1557 GATGCCATCTCAAGAGACGGTTCACCTGATACAAATCCCATCCATGCTGATTCATC 1616
      |||||
Db 548 GCTGCCGTGTGAGTCTCAACCGAGAGCTGTACCCAGTGTGATGCTGTGGTCTTCACC 607
OY 1617 ACCGACATCATGTTGGATCTGACAAATTTACCTACGACAGCTAGGCCACCTCCACA 1676
      |||||
Db 608 ACCGTAGCTGCAACCGGCAATCTCTCTACCCCTGGACAGGAGGCAACGAGACG 667
OY 1677 AATGATTTGATGAAATTTGAAATCACCACATCCACATCCCAAAAAGAGTGCATTGAGC 1736
      |||||
Db 668 CTTGGCTCTGGGCTCTCCATGGAATCGCCAGATATACCATGATGTCATCGCTCCGGG 727
OY 1737 ACTTGTAACTGACTGACTGACTGACGCGGTGATTCAGATATCCAAAGTCCATTAGGCT 1796
      |||||
Db 728 GCATCTTCAACTGGGTGGGTGAGTATGGGGTATTCAGATATCTTTGATCCCTACGGTC 787
OY 1797 TCTTGACGGTAGACACAAATCCCTTGTGTTGATGTCGCAAGCAAGAAATTTGCT 1856
      |||||
Db 788 GCTTGGAGCCCTCTCTCTGGGCC---CACATCCCACTACGCGCAAAAGCTGAGTGC 844
OY 1857 GCTGGCTTGTGATGATGAGAACCCCTGAGCATCCAGATCAAGATTTACAAAGAGCTAA 1916
      |||||
Db 845 CCGTGGGTGATGACAAATTTCCAGAGGCGGACGAGAGCTGACAAAGCTGACGCGAGCTGG 904
OY 1917 GCAAAAGCATTTGAATCCATCTACGAGGCAAGCATTTGGAGATATGTCAATGATATAA 1976
      |||||
Db 905 CCCCTCATCTGACAGGTGATGTGTCGTCGCGCCAGCGAGCGGCCCTATGCGCTAAT 964
OY 1977 ATTGATTCCTACATATCTGCTGTAATTTATGTTTCCTTTGAAAAATTCATTCACA 2036
      |||||
Db 965 GTCTGCTGCCACTTTGGCCCGGATACCGCTTCTACCTGGCTTTGAGAACTACAGCATC 1024
OY 2037 AGATTTACATCAGGAAAAAGCTAT---ACAATGCTTTTCTGGCTGGCTGTACCTGTG 2093
      |||||
Db 1025 GGGACTACATCAGTGAAGTGTGGCGCAATGCTCGGCGGCTGTCTGTACCCCTGG 1084
OY 2094 TTCTGGACCATCTAGGAAAACTAGTAATATATTCACAGATTCATTCAATTCATG 2153
      |||||
Db 1085 CGGTGGACCTCTCTGGGCGACCTACGAGAGCTTTGTGGACCAAGATGCTTTGTACAG 1144
OY 2154 TGGAGATTTAATCTCTCCAGTAGAGCTAGC 2184
      |||||
Db 1145 TGGACGACTTCAGCTCTGCCGCTGAACTGCG 1175
```

RESULT 3

```
US-07-914-281-13
: Sequence 13, Application US/07914281
: Patent No. 5324663
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B.
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
: TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
```

```

? ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
? ADDRESS: P.C.
? STREET: 1755 Jefferson Davis Highway, Fourth Floor
? CITY: Arlington
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/914,281
? FILING DATE: 19920720
? CLASSIFICATION: 530
? ATTORNEY/AGENT INFORMATION:
? NAME: Lavalleye, Jean-Paul M. P.
? REGISTRATION NUMBER: 31,451
? REFERENCE/DOCKET NUMBER: 2363-060-55
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703)521-4500
? TELEFAX: (703)486-2347
? TELEX: 248855 OPAT UR
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1654 base pairs
? TYPE: NUCLEIC ACID
? STRANDEDNESS: unknown
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? US-07-914-281-13

Query Match          4.0%; Score 113.4; DB 1; Length 1654;
Best Local Similarity 51.2%; Pred. No. 1.1e-15;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

QY 1484 ATCTGCTGTGGTGGTGGCCATTGGGCGAGACCTTTGACCTTACATCTGCCAAGCAATG 1543
DB 318 ATCTGCTGTGGTGGTGGCCATTGGGCGAGACCTTTGACCTTACATCTGCCAAGCAATG 377
QY 1544 TT---CAACATCCAGAGTGGCCATCTCACAAGCGACCGCTTCACTGTACCAAAATCCCAT 1600
DB 378 GTGCTGGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
QY 1601 GCAGTCTGTGATCCATCAGCAGACATCAGTGGGATCTGA---CAATTTACCTCAGCAA 1657
DB 438 GCGGTATGCTGTCACACCGACGAGTCAATGACACCCCAAGTCCCAAGCTCCCAAGCTCC 497
QY 1658 GCTAGGCCACCTTCCAGAAATGATTTGGATGATGATTTGGAATCACAACCTCAGACTCC 1717
DB 498 CCGAGGCGGCGAGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 557
QY 1718 CAAAAGAGTGGCATTTAGACACTTGTTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1777
DB 558 CAACTTAAGAGCCATGAGCAGATCTCAATCTCACCATCTCCACGCGACTCCGAC 617
QY 1778 ATCCAGTGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1828
DB 618 ATCTTACGCGCTTACGCGCTGCTGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
QY 1829 GAAGTCCCAAGCAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1888
DB 678 AACCTCTCGGCCCAAGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
QY 1889 GCGAGAGTCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1948
DB 738 GCGAGGCTGCGCTTACCTCAGACGAGCCCATCTCAAGGTGACGCTGACG---A 794
QY 1949 GCATTTGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2008
DB 795 CGCTCCACAGAGCCCTGCCCGCCAGGACCATGATGAGACGCTGTCCGGGTACAAAGTTC 854

```

```

QY 2009 TATCTTCTCTTGAATAATTCACAAAGATTTACATCAGGAAAGCTAT---ACAAT 2065
DB 855 TATCTGCGCTTCGAGAACTCTTGACACCGGACTACCGAGAACTGTGGAGAAC 914
QY 2066 GCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2125
DB 915 GCCCTGAGGCGCTGGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974
QY 2126 TATATTCACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2185
DB 975 TTCTGTCACCGACCGAGCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1034
QY 2186 AAGTATCTGAGGAAGTTCGACAAATTAATGATTAATGATTAATGATTAATGATTAAT 2245
DB 1035 CGGTACTGAGGAGGCTGAGACAGACACGCGGCTACTGAGCTACTTCTGCTGCGG 1094
QY 2246 AAG 2248
DB 1095 GAG 1097

RESULT 4
US-08-393-246-13
? Sequence 13, Application US/08393246
? Patent No. 5595900
? GENERAL INFORMATION:
? APPLICANT: LOWE, JOHN B.
? TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
? TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
? TITLE OF INVENTION: GLYCOPOLYDIPS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
? STREET: 1755 Jefferson Davis Highway, Fourth Floor
? CITY: Arlington
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/393,246
? FILING DATE:
? CLASSIFICATION: 530
? PRIOR APPLICATION NUMBER:
? APPLICATION NUMBER: US 08/220,433
? FILING DATE: 30-MAR-1994
? APPLICATION NUMBER: US 07/914,281
? FILING DATE: 20-JUL-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Lavalleye, Jean-Paul M. P.
? REGISTRATION NUMBER: 31,451
? REFERENCE/DOCKET NUMBER: 2363-060-55
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703)521-4500
? TELEFAX: (703)486-2347
? TELEX: 248855 OPAT UR
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1654 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: unknown
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? US-08-393-246-13

Query Match          4.0%; Score 113.4; DB 1; Length 1654;
Best Local Similarity 51.2%; Pred. No. 1.1e-15;

```

Tue Oct 8 10:18:24 2002

us-09-744-748-4.rn

Page 4

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

OY	1484	ATCTGGTGGGGTGTGGGCATTTTGGGGAGACCTTTGACCACTTAACTGTCGACGAAAG	1543
Db	318	ATCTGCTGTGGACGTGGCCCTTTTAAACAACCACTAGCTGTGCCCGCTGTGAGATG	377
OY	1544	TT---CAACATCCAAAGATGCGCATCTTCACAGGACCGTTCACTGTACACAATCCCAT	1600
Db	378	GTGCGCTGGCAGCGCTGACTGACATATCACTGCGCAGCAAGGTGTATCCACAGGAGAC	437
OY	1601	GCAGTTGATTCATCTACCGAGACATCACTTTGGGATGTGA---CAAAATTACTCAGCAA	1657
Db	438	GGCGTCATCTGTGACACCGAGAGATGTATGTACAAACCCAGTCCCACTCCCAACGCTCC	497
OY	1658	GGTAGGCAACCCCTTCCAAATATGATTTGGATGAAATTGGAAATCAACAATCTCACTGCC	1717
Db	498	CCGAGGCGGAGAGGGGACGACGATGATCTGGTTTCAGCATGAGATGCCCAAGCCATGCTGG	557
OY	1718	CAAAAGATGCGATTGACACACTTGTTTAACTGACTGTGACTTACCGCGGTATTCAGAT	1777
Db	558	CAGCTGAAACCATGAGAGGATATCTCAATCTCAACATGTCTACCGAGGAGTCTCGAC	617
OY	1778	ATCCAAATGCTTATGGCT-----TCTTGAAGGATGAAGCAAAATCCCTTGCTGTTT	1828
Db	618	ATCTTCAACGCCCTCAGCGCTGGCTGGAGCGCTGTGTCGGGACAGCTGCGCCACCAACGCGTC	677
OY	1829	GAAAGGCCAAGCAAAAGAAATTTGGTGTCTGTGGGTGTGAGTAACTGCAACCCAGACAT	1888
Db	678	AACTCTCGGCCAAGACCGAGACTGTGTGCGCTGGGCAAGTCCCACTGGGGGCAAACTCC	737
OY	1889	GCAGAGTCAAGTATTACATAGAGCTTAAGCAAAAGCATTAATTCATATCTTACGGGCAA	1948
Db	738	GGCAGAGGGGCGTACTACCAAGGCTGCGAGGCCCATCTCAAGGTGAGCTGTAGG---A	794
OY	1949	GCATTTGGAGAATATGTCAATGATTAATAAATTTGATCTCTACATATCTGTGTTGAATT	2008
Db	795	CGCTCCCAACAAGCCCTCGCCCAAGGACCAATGATGAGAGCGTGTCCCGGTACAAAGTTTC	854
OY	2009	TATCTCTTCCTTTAAAAATTCAATCCACAGAGATTATCAATCCAGGAAAAAGCTAT---ACAAT	2065
Db	855	TATCTGGCCTTCAGAAATCTCTTGCACGCCGACTACATCAAGAAAGCTGTGGAGAAC	914
OY	2066	GCTTTTGGCTGGCTCTGTACTGTTTGTCTGGGACCATCTTAGGGAAAATATGAAAT	2125
Db	915	GGCTCGAGAGCGTGGAGCGGTCCCGTGGTGGTGGGCCCCAGCAAGACACTTACGAGAGG	974
OY	2126	TATATTCCACAGATTCATTTATTCATGTGGAAGATTATATACCTCCCAAGTACATGAGA	2185
Db	975	TTCGTGCGACCCGACCCCTTCATCAAGTGGAGCACTTCAAGAGCCCCCAAGACGTGGCC	1034
OY	2186	AAGTATCGAAGAAAGTGCACAAAACAAATAAATTAATCTATGTAATCTTAACTAGAG	2245
Db	1035	CGGTAACTCGAAGAGCTGACAAAGAACACGCCCGCTACCTGAGCTACTTGTGCTGGCG	1094
OY	2246	AAG 2248	
Db	1095	GAG 1097	

RESULT 5
US-08-525-058A-13
; Sequence 13, Application US/08525058A

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS OF OLIGOSACCHARIDE STRUCTURES ON GLYC

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

NUMBER OF SEQUENCES: 2

ADDRESSEE: ORION SPRAY MCGILL AND MATHEW MCNEILSTADT B C

STREET: 1755 Jefferson Davis Highway, Fourth Floor,

TOOT 7 10 2000 / 1999

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,058A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPMT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs

MOLECULE TYPE: DNA (genomic)
US-08-525-058A-13

Query Match 4.08; Score 113.4; DB 1; Length 1654;
Post Local simlari+... E1 39... prod no 1 12-15.

Matches	401:	Conservative	0:	Mismatches	361:	Indels	21:	Gaps	5:
Best Local Similarity	51.26;	Pred. NO.	1.1e-15;						

OY	1484	ATTTCGGTGGGAGTGGGACATTGGGACAGACCTTGACCTTACATCTCCGACAGCAT	1543
Db	318	ATCTGGCTGGGACGCGGGCTTTTAAACAACCCANAGCTGCCCGGGCTCGAGAGT	377
OY	1544	TT---CAACATCCAGATAGTGCATCTACAGAGACGCTTCACTGTACACAAATCCCAT	1600
Db	378	GTGCTGGACAGCGCTACTGCATCATCTGCCGACCGCAGAGTGTATCCACAGGACGAC	437
OY	1601	GCAATTCTGATTCATCCAGACGAGCATTCAGTGGATGTGA--CAAAATTACTCGACAA	1657
Db	438	GGGGTCACTCGTGCACACCCAGAGAGTCAATGCAACCCAGTCCCGAGCTCCACAGTCC	497
OY	1658	GCTAGCGCCACCTTCCAGAAATGGATTGGATGAATTTGGAAATACACAATCACTCCG	1717
Db	498	CCGAGCGCGCAGGGGAGGAGCATGTCGTTCACATGAGAGACCCCAAGCCACTCGTG	557
OY	1718	CAAAAGATGGACATGAGCACTTTTAAACGTGACTGATCCGCGCGATGATCGAT	1777
Db	558	CAGCTGAAGACATGACGAGTACTTCATCTCACATGTCTCTACCGACGACTCCGAC	617
OY	1778	ATCCAAATGCCCTTATGCT-----TCTTGACGGTAAGCAACAATCCCTTGCTTT	1828
Db	618	ATCTTCAGCGCCGACGGCTGGCTGAGACGCTGTCCGGACGCTCCACACACCGCTC	677
OY	1829	GAGTGCACAGCAAAAGAAATTTGCTGTGCTGGTGTGGATGTAATGCAACCTCGACAT	1888
Db	678	AACCTCTGGGCAAGACACGACCTGTGGCTGGGAGGTGCATCAACTGGGGGCAAACTGC	737
OY	1889	GCCACAGTCAAGATTATCAATAGACTTAAGCAAAACATTTGAATTCATACCTACGGGCA	1948
Db	738	GCGAGGTTGGGCTACTACGAGGCTCGAGGCCATCTCAAGGTGGAGGTGTACGG---A	794
OY	1949	GCAATTGAGAAATATCTCAATGATTAATAATTTGATTTCTTACATATCTGTTGAATTT	2008
Db	795	CGCTCCCAACAGCCCTGCTCCCGAGGAAACCATGATGAGAGCGCTGTCCGGATCAAACTTC	854
OY	2009	TATCTTCTCTTTGAAAATTCATCCACAGAGATTCATCAACGAGAAAGCTAT---ACAT	2065
Db	855	TATCTGGCTCTGACAAAGATCTCTGTACACCCGAGTCAATACACAGAAAGCTGTGGAGAC	914

Tue Oct 8 10:18:24 2002

CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC DOS/MS-DOS
 CURRENT: Patent Release #1.0, Version #1.25
 APPLICATION DATA:
 FILING DATE: US/09/042.531
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE: US 08/220,433
 FILING DATE: 30-MAR-1994
 APPLICATION NUMBER: 20-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Luc
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: P.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703/3486-2347
 TELEFAX: 703/3486-2347
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1654 base pairs
 TYPE: nucleic acid
 STANDARDS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-09-042-531-13

[illegible]

us-09-744-748-4.rn1

[illegible]

RESULT 8
 US-07-914-281-12
 Sequence 12
 Patent No. 5124663
 GENERAL INFORMATION
 APPLICANT: LOME, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES IN GLYCOPROTEINS
 TITLE OF INVENTION: GLYCOLIPIDS OR AS FREE MOLECULES, AND FOR THE ISOLATION
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 19920720
 FILING DATE: 07/914, 281
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31451
 REFERENCE/DOCKET NUMBER: 363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-07-914-281-12
 Query Match
 Best Local Similarity 4.0%; Score 112.2; DR 1;
 Matches 364; Conservative 0; Mismatches 325; Indels 18; Gaps 0;

```

QY 1559 TGCATCTCACAACGACCGCTTACATGATACAAATCCCATGCTTGTATCATCAC 1618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TGCACATCTACGCGCGCCGCAAGGTGTACCAAGAGCAACAGGTCATCGTACACAC 330
QY 1619 CGAGACATCAGTTGGGA---TCTGACAAATTTTACCTCAGCAAGCTAGCCACTTCCAG 1675
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TGGGATATCATGTGTCACACCTTAAGTACAGGCTCCACCTTCCCGCGGCGGAGG 390
QY 1676 AATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CGCTGATCTGCTTCAACTTGGAGCCACCCCTTAAGTCCAGACCTGGAAGCCCTGGAG 450
QY 1736 CACTTGTATTAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AGATCTTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 510
QY 1796 TTCTTACGCTAAGCAACAATCCCTTCTG-----GTTTGAAGTGGCAAGCAAGAG 1846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 TGGCTGAGCGCGTGTCCGCGCCAGCTGCGCCACCCGCTCAACCTCTCGGCGCAAGACC 570
QY 1847 AATTTGCTGTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1906
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GACCTGTGCTGCTGGCGGCTGTCTCAACTGGAAGCCGACTCAGCCAGGCTGCTACTAC 630
QY 1907 AATGACCTAAGCAAAAGCATTTGAATCCATACCTACGCGGCAAGCATTTGAGATATGTC 1966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAGAGCTGAGGCTCATCTCATAGTGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 690
QY 1967 AATGATTAATAATTTGATTTCTTCAATCTGCTTGAATTTTATCTTCTTCTTGAAT 2026
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 AAGGGAGCA---TGATGAGAGAGCTGTCCCGTACAACTTCACTGCTGCTGCGAAGAC 747
QY 2027 TCAATCCAAAGATATACATCAAGCAAAAGCTAT---ACAATGCTTTTCTGCTGCTCT 2083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 TCCCTTCCACCCGCTACATCACTCAAGCAAGAGCTGTGAGAGAGCCCTGAGAGCTGAGCC 807
QY 2084 GTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 867
QY 2144 TTCTTCAATGTGGAATATATATATCTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAG 2203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 TTCTTCAATGTGGAATATATATATCTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAG 927
QY 2204 GACAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 GACAAAGACACGCGCGCTACCTGAGCTACTTTCGCTGGCGGAG 972

```

```

RESULT 9
US-08-393-246-12
: Sequence 12, Application US/08393246
: Patent No. 5595900
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B.
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
: TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
: ADDRESS: P.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/393,246
: FILING DATE:
: CLASSIFICATION: 530
: PRIORITY INFORMATION:
: APPLICATION NUMBER: US 08/220,433
: FILING DATE: 30-MAR-1994
: APPLICATION NUMBER: US 07/914,281
: FILING DATE: 20-JUL-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Lavelleye, Jean-Paul M. P.
: REGISTRATION NUMBER: 31,451
: REFERENCE/DOCKET NUMBER: 2363-060-55
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)521-4500
: TELEFAX: (703)486-2347
: TELEX: 24885 OPAR UR
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1086 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: US-08-393-246-12

```

Query Match 4.0%; Score 112.2; DB 15; Length 1086;

Best Local Similarity 51.6%; Pred.No. 1,8e-15; Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

QY 1559 TGCATCTCACAACGACCGCTTACATGATACAAATCCCATGCTTGTATCATCAC 1618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TGCACATCTACGCGCGCCGCAAGGTGTACCAAGAGCAACAGGTCATCGTACACAC 330
QY 1619 CGAGACATCAGTTGGGA---TCTGACAAATTTTACCTCAGCAAGCTAGCCACTTCCAG 1675
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TGGGATATCATGTGTCACACCTTAAGTACAGGCTCCACCTTCCCGCGGCGGAGG 390
QY 1676 AATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CGCTGATCTGCTTCAACTTGGAGCCACCCCTTAAGTCCAGACCTGGAAGCCCTGGAG 450
QY 1736 CACTTGTATTAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AGATCTTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 510
QY 1796 TTCTTACGCTAAGCAACAATCCCTTCTG-----GTTTGAAGTGGCAAGCAAGAG 1846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 TGGCTGAGCGCGTGTCCGCGCCAGCTGCGCCACCCGCTCAACCTCTCGGCGCAAGACC 570
QY 1847 AATTTGCTGTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1906
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GACCTGTGCTGCTGGCGGCTGTCTCAACTGGAAGCCGACTCAGCCAGGCTGCTACTAC 630
QY 1907 AATGACCTAAGCAAAAGCATTTGAATCCATACCTACGCGGCAAGCATTTGAGATATGTC 1966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAGAGCTGAGGCTCATCTCATAGTGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 690
QY 1967 AATGATTAATAATTTGATTTCTTCAATCTGCTTGAATTTTATCTTCTTCTTGAAT 2026
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 AAGGGAGCA---TGATGAGAGAGCTGTCCCGTACAACTTCACTGCTGCTGCGAAGAC 747
QY 2027 TCAATCCAAAGATATACATCAAGCAAAAGCTAT---ACAATGCTTTTCTGCTGCTCT 2083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 TCCCTTCCACCCGCTACATCACTCAAGCAAGAGCTGTGAGAGAGCCCTGAGAGCTGAGCC 807
QY 2084 GTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 867
QY 2144 TTCTTCAATGTGGAATATATATATCTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAG 2203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

us-09-744-748-4.rni

Page 8

[illegible]

TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-696-731-12

Query Match 4.0%; Score 112.2; DB 2; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 1.8e-15;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

1559 TGGCATCTCAACAGGACGGTTCACGTACACAAATCCCATCAGTTCATCATCAC 1618
 271 TGGCATCTCACTCGACGCGACGAGGTATCCACAGCAGACGGTCACTGACAC 330
 1619 CGAGACATCAGTTGGGA---TCTGACAAATTTACCTCAGCAGCTAGGCCACCTTCCAG 1675
 331 TGGGATATCATGTCCAAACCTTAAGTACAGCCTCCACCTTCCCGAGGCGGAGGAG 390
 1676 AATGATTTGATGATGATTTGGAATCAGACACACACCTCCCAAAAGAGTGCATTGAG 1735
 391 CGCTGATCTGTTCAACTTGGAGCCACCCCTTAACCTGCGACGACCTGGAAGCGCTGAG 450
 1736 CACTTGTTAACCTGACTGACTGACTACCGCGCGATTCAGATATCCAAAGTCCCTTATGGC 1795
 451 AGATACCTCAATCTCAACATGTCCTACCGGAGGACTCCGACATCTTCAAGCCCTACGGC 510
 1796 TTTCTACGCTTAAGCAACAATCCCTTCGT-----GTTTGAAGTCCCAAGCAAGAG 1846
 511 TGGCTGAGCGCTGTCGCGCCAGCCTGCCACCCGCTCAACCTCTCGGCAAGACC 570
 1847 AATTTGTTGCTGGTGTGATGATGATGAGTGAACCTGAGATCCGAGAGTCAAGTATTAC 1906
 571 GACTGTGGCTGGCGGCTGCTCAACTGGAAGCCGAGCTCAACGAGGCTGCTTACTAC 630
 1907 AATGACTAAGCAAAAGCATTTGAATTCATACCTACGAGGCAAGCATTTGGAAATATGTC 1966
 631 CAGAGCTGAGGCTCATCTCAAGGTGAGGTGAGGTACGAGCGCTCCCAAGCCCTGCCC 690
 1967 AATGATAAAAATTTGATTCCTACCAATCTGCTTGAATTTTATCTTCCCTTGAATAT 2026
 691 AAGGAGACCA---TGATGAGAGCGTGTCCGGGTACAAAGTTTACTGCGCTTCGAGAAC 747
 2027 TCAATCCACAAGATTAATCAAGGAAAAGCTAT---ACAATCTTTTCTGGCTGCTCT 2083
 748 TCTTTCACCCCGACCTACATCAACGAGAGCTGTGAGGAGAACCCCTGAGGCTGAGGC 807
 2084 GATCTGTTGTTCTTGGACCATCTAGGAAACTATGAGAAATTAATTCAGCAGATTGA 2143
 808 GTTCCCGTGTGTGTGGGCCCCCAGACAGAACTACGAGAGGTCTCTGACACCGGAGGC 867
 2144 TTTCAATTCATGGAAGATTATTAATCTCCAGTGAAGCTGACCAAGATATCTGAAGAGTC 2203
 868 TTTATCATGAGTGAAGACTTCCAGAGCCCAAGAGACTGAGCCGAGTACTCGAGAGCTG 927
 2204 GACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2248
 928 GACAAAGACCAAGCCCGCTACCTGAGCTTCTGCTGGCGGAG 972

RESULT 12

US-09-042-531-12
 Sequence 12, Application US/09042531

GENERAL INFORMATION:
 APPLICANT: LOME, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/042,531
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:

APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994

APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:
 NAME: Lavalley, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-09-042-531-12

Query Match 4.0%; Score 112.2; DB 4; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 1.8e-15;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

1559 TGGCATCTCAACAGGACGGTTCACGTACACAAATCCCATCAGTTCATCATCAC 1618
 271 TGGCATCTCACTCGACGCGACGAGGTATCCACAGCAGACGGTCACTGACAC 330
 1619 CGAGACATCAGTTGGGA---TCTGACAAATTTACCTCAGCAGCTAGGCCACCTTCCAG 1675
 331 TGGGATATCATGTCCAAACCTTAAGTACAGCCTCCACCTTCCCGAGGCGGAGGAG 390
 1676 AATGATTTGATGATGATTTGGAATCAGACACACCTCCCAAAAGAGTGCATTGAG 1735
 391 CGCTGATCTGTTCAACTTGGAGCCACCCCTTAACCTGCGACGACCTGGAAGCGCTGAG 450
 1736 CACTTGTTAACCTGACTGACTGACTACCGCGCGATTCAGATATCCAAAGTCCCTTATGGC 1795
 451 AGATACCTCAATCTCAACATGTCCTACCGGAGGACTCCGACATCTTCAAGCCCTACGCG 510
 1796 TTTCTACGCTTAAGCAACAATCCCTTCGT-----GTTTGAAGTCCCAAGCAAGAG 1846
 511 TGGCTGAGCGCTGTCGCGCCAGCCTGCCACCCGCTCAACCTCTCGGCAAGACC 570
 1847 AATTTGTTGCTGGTGTGATGATGAGTGAACCTGAGATCCGAGAGTCAAGTATTAC 1906
 571 GACTGTGGCTGGCGGCTGCTCAACTGGAAGCCGAGTCAAGGAGTGGGCTGACTAC 630
 1907 AATGACTAAGCAAAAGCATTTGAATTCATACCTACGAGGCAAGCATTTGGAAATATGTC 1966
 631 CAGAGCTGAGGCTCATCTCAAGGTGAGGTGAGGTACGAGCGCTCCCAAGCCCTGCCC 690

Tue Oct 8 10:18:24 2002

us-09-744-748-4.rn1

Page 12

```
Db 583 TGCTGGAGCGCGTGTGTCGGCCAGCCCTGCCCACCCACCCGCTCAACCTCTGGCCNAGACC 642
QY 1847 AAATTGGTGTGCTGGGTTGTGAGTAACCTGGAACCCCTGAGCATTGCCAGAGTCAAGTATTAC 1906
Db 643 GAGCTGGTGCTGGGGGTGTCTCAACTGGAAGCCGACTCAGCCAGGGTGGCTACTAC 702
QY 1907 AATGAGCTAAGCAAAAGCAATGAAATCCATACCTACGGGCAGACATTGGAGAAATATGTC 1966
Db 703 CAGAGCCTGCAGGCTCATCTCAAGGTGAGACGTGTACGGACGTCGCCACAAGCCCTGCCC 762
QY 1967 AATGATMAAAATTTGATTCTCTACCATATCTGCTGTAAATTTTATCTTCCTTTGAAT 2026
Db 763 AAGGGGACCA--TGATGAGACGCTGTCCGGTACAAAGTTCTACCTGGCCTTCGAGAAC 819
QY 2027 TCAATCCACAAGGATTACATCAGGAAAAGCTAT--ACAATGCTTTTCTGGCTGGCTCT 2083
Db 820 TCCTTGCACCCCGACTACATCACCGAGAGAGCTGTGAGGAAACGCCCTGGAGGCCCTGGGCC 879
QY 2084 GTACCTGTGTCTCTGGACCACATCTAGGGAAAACATAGAGAAATATATATTCACGACATTC 2143
Db 880 GTGCCCGTGTGCTGGGCCCCACAGAGCAAGCACTACGAGAGTTCTCCGCCACCCGACGCC 939
QY 2144 TTCATTGATGTGGAGATTTATTAACCTCTCCAGTACGAGCTAGCAAAAGTATCTGAAGAAATC 2203
Db 940 TTCATCCAGCTGAGCACTCCAGAGCCCCCAGAGACCTGGCCCGGTACCTGCAGAGAGCTG 999
QY 2204 GACAAAAACAATAAGTTATACCTTAGTTACTTAACTGGAGGAAG 2248
Db 1000 GACAAGGACCAAGCCGCTACGTGAGCTACTTTCGCTGGCGGGAG 1044
```

Search completed: October 6, 2002, 07:21:20
Job time : 103.579 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:39:45 ; Search time 1353.36 Seconds
(without alignments)
16699.717 Million cell updates/sec

Title: US-09-744-748-4_COPY_1289_2368

Perfect score: 1080

Sequence: 1 atgacatcaacatccaagg.....agaaatggttggaattaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Database :

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

GenEmbl:*
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	1080	100.0	3019	9	AB023021	AB023021 Homo sapi
2	1078.4	99.9	80247	9	AL512406	AL512406 Human DNA
3	1075.2	99.6	2501	9	HS238701	AJ238701 Homo sapi
4	931.2	86.2	2139	10	AB015426	AB015426 Mus muscu
5	929.6	86.1	1128	10	AF345993	AF345993 Rattus no
6	928	85.9	2156	10	AB049819	AB049819 Rattus no
7	916.8	84.9	1705	10	AF230460	AF230460 Cricetulu
8	783.6	72.6	1707	5	AB035906	AB035906 Gallus ga
9	676.8	62.7	1080	5	AB035905	AB035905 Xenopus l
10	639	59.2	71396	2	AC100674	AC100674 Mus muscu
11	277.8	25.7	1567	5	AB023627	AB023627 Danio rer
12	234.6	21.7	1832	5	AB023628	AB023628 Danio rer
13	222.6	20.6	1092	5	AB035907	AB035907 Xenopus l
14	220.2	20.4	1264	5	AF288369	AF288369 Gallus ga
15	176.2	16.3	2657	5	GGU73678	U73678 Gallus gall
16	145.2	13.4	1429	10	CG878737	AF090450 Cricetulu
17	145.2	13.4	1462	10	AF090450	AF090449 Cricetulu
18	143.6	13.3	1718	10	AF090449	AB035908 Gallus ga
19	137.2	12.7	1044	5	AB035908	X87810 B.taurus al
20	132	12.2	2293	4	BTA134FTG	AJ132772 Bos tauru
21	132	12.2	12607	2	BTA132772	AC100674 Mus muscu
22	128	11.9	71396	2	AC100674	AB039126 Mus muscu
23	127.6	11.8	1158	10	AB039126	AF016889 Schistoso
24	127	11.8	1487	3	AF016889	AF016889 Sequence
25	127	11.8	1814	6	AR028686	U45980 Mus muscu
26	127	11.8	3594	10	MMU045980	AC091284 Mus muscu
27	127	11.8	117993	2	AC091284	AJ132774 Bos tauru
28	125.6	11.6	1613	4	BTA132776	AJ132775 Bos tauru
29	125.6	11.6	1619	4	BTA132775	AJ132776 Bos tauru
30	125.6	11.6	1644	4	BTA132774	AJ132773 Bos tauru
31	125.6	11.6	1873	9	AF345881	AF345881 Macaca mu
32	124.6	11.5	1119	9	AF345881	AB039131 Mus muscu
33	122.8	11.4	1170	10	AB039127	U33457 Mus muscu
34	122.8	11.4	1170	10	AB039131	AF345884 Saimiri s
35	122.8	11.4	1615	10	MMU03457	AB039124 Mus muscu
36	122	11.3	1113	9	AF345884	AB039125 Mus muscu
37	121.2	11.2	1170	10	AB039124	AB039128 Mus muscu
38	121.2	11.2	1170	10	AB039125	AB039129 Mus muscu
39	121.2	11.2	1170	10	AB039128	AB039130 Mus muscu
40	121.2	11.2	1170	10	AB039129	AB039132 Mus muscu
41	121.2	11.2	1170	10	AB039130	AF345883 Macaca mu
42	121.2	11.2	1170	10	AB039132	AB039133 Mus spici
43	119.8	11.1	1119	9	AF345883	AB039133 Mus spici
44	118.4	11.0	1170	10	AB039133	AB051859 Pongo pyg
45	118.2	10.9	1080	9	AB051859	

ALIGNMENTS

RESULT 1
AB023021
LOCUS Homo sapiens FUT9 mRNA for alpha-1,3-fucosyltransferase IX, complete cds.
DEFINITION
ACCESSION AB023021
VERSION AB023021.1 GI:5139692
KEYWORDS alpha-1,3-fucosyltransferase IX; FUT9.
SOURCE Homo sapiens CDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (sites)
REFERENCE
AUTHORS Kaneko,M., Kudo,T., Iwasaki,H., Ikehara,Y., Nishihara,S., Nakagawa,S., Sasaki,K., Shitina,T., Inoko,H., Saitou,N. and Narimatsu,H.
TITLE Alpha 1,3-fucosyltransferase IX (Fuc-TIX) is very highly conserved between human and mouse; molecular cloning, characterization and tissue distribution of human Fuc-TIX
JOURNAL FEBS Lett. 453, 237-242 (1999)
REFERENCE
AUTHORS Kaneko,M., Kudo,T. and Narimatsu,H.
TITLE Direct Submission

JOURNAL Submitted (29-JAN-1999) Mika Kaneko, Institute of Life Science,
Soka University, Division of Cell Biology, Soka University, 1-236
Tangi-cho, Hachioji, Tokyo 192-8577, Japan
(E-mail:mika@soka.ac.jp, Tel:81-426-91-2495(ex.5132),
Fax:81-426-91-9315)

FEATURES

Location/Qualifiers

source

1..3019

/organism="Homo sapiens"

/db_xref="taxon:9606"

gene

295..1374

/gene="Fur9"

CDS

295..1374

/gene="Fur9"

/codon_start=1

/product="alpha-1,3-fucosyltransferase IX"

/protein_id="BA81685.1"

/db_xref="GI:5139693"

/translation="MTSTSKGILRPLIVCIILGCFMACLIYIKPTNSWTFSPWESA
SSVLKMNFEPTKDYFENETTLVWVPGQGFDTLSCAMENIGCHLTDRSLYRK
SHAVLIHHRDISDLNLPQARPPQKIMNLESPPTPKSGIEHLFLTLTYR
DSDIQVYGLFSTNFEVEFVPSKELVQWVSNMNPENHAKYKYNELKSIEHTY
QGAQEVYNDKNIIPITISACKFYLSFENSIHKDYIEKLYLAGSVYVLAGSPREN
YENIIPADSEFHEVDYNSPELAKYLEKVDKNNKLYLSYFNMRKDFVNLPRMESH
CLACDHYRHOEKSQGNLEKFWN"

BASE COUNT 937 a 594 c 541 g 947 t

ORIGIN

Query Match 100.0%; Score 1080; DB 9; Length 3019;
Best Local Similarity 100.0%; Pred. No. 1.2e-268;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACATCAACATCCAAAGAAATTCCTGCCATTTTAAATGCTGCATTAATCTGGCC 60
DB 295 ATGACATCAACATCCAAAGAAATTCCTGCCATTTTAAATGCTGCATTAATCTGGCC 354
QY 61 TGTTCATGCGATGCTTCCTCATTTATCAATCAACCACTGAGATCTTCAGTCCA 120
DB 355 TGTTCATGCGATGCTTCCTCATTTATCAATCAACCACTGAGATCTTCAGTCCA 414
QY 121 ATGGAATCAGCAGCTCTGTGCTGATAAATGAAATCTTTCCACCAAACTGATTT 180
DB 415 ATGGAATCAGCAGCTCTGTGCTGATAAATGAAATCTTTCCACCAAACTGATTT 474
QY 181 TTTAATGAAGAACTACTATTCCTGTGGGTGGGCGCATTTGGCGAGACCTTTGACCTTAA 240
DB 475 TTTAATGAAGAACTACTATTCCTGTGGGTGGGCGCATTTGGCGAGACCTTTGACCTTAA 534
QY 241 TCTGCGCAAGCAATGTTCAATCAATCAAGATGCGCATCTCACACGACCGCTTACTGTAC 300
DB 535 TCTGCGCAAGCAATGTTCAATCAATCAAGATGCGCATCTCACACGACCGCTTACTGTAC 594
QY 301 AACAAATCCCATGAGTTCTGATCCATCCAGAGACATCAGTTGGATCTGACAAATTTA 360
DB 595 AACAAATCCCATGAGTTCTGATCCATCCAGAGACATCAGTTGGATCTGACAAATTTA 654
QY 361 CCTCAGCAAGTAGGCGCACCTTCAGAAATGATTTGATGAATTTGAATCACCACACT 420
DB 655 CCTCAGCAAGTAGGCGCACCTTCAGAAATGATTTGATGAATTTGAATCACCACACT 714
QY 421 CACACTCCCAAAAGAGAGGATGAGACATGTTTAACCTGACTGACTGACTGACTGACTGACT 480
DB 715 CACACTCCCAAAAGAGAGGATGAGACATGTTTAACCTGACTGACTGACTGACTGACTGACT 774
QY 481 GATTCAGATATTCAGATGCTTATGCTTCTTGACGGTAAAGCACAATCCCTTGCTGTT 540
DB 775 GATTCAGATATTCAGATGCTTATGCTTCTTGACGGTAAAGCACAATCCCTTGCTGTT 834
QY 541 GAAGTGCACAGCAAGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 835 GAAGTGCACAGCAAGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 601 GCGAGAGCAAGTATTAATGAGCTAAGCAAAAGCATTTGAATTCATCTACGGGCA 660

DB 895 GCCAGAGTCAGTATTTACATAGCTAAGCAAAAGCATTTGAATTCATACCTACGGCGCA 954
QY 661 GCATTTGAGAAATATGTCATGATTAATAATTTGATTCCTACATATGCTGTTAAATTT 720
DB 955 GCATTTGAGAAATATGTCATGATTAATAATTTGATTCCTACATATGCTGTTAAATTT 1014
QY 721 TATCTTCCCTTTGAAATTCATCCACAGGATTATCATCAGCAAGAAAGCTATTCATGCT 780
DB 1015 TATCTTCCCTTTGAAATTCATCCACAGGATTATCATCAGCAAGAAAGCTATTCATGCT 1074
QY 781 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 1075 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
QY 841 ATTCAGCAGATTCATTCATGCTGAGAAATTAATTAATCTCCAGTAGCTAGCAAG 900
DB 1135 ATTCAGCAGATTCATTCATGCTGAGAAATTAATTAATCTCCAGTAGCTAGCAAG 1194
QY 901 TATCTGAGAGAGTGCACAAAACAAATTAATTAATCTTACTTAACTGAGAGAG 960
DB 1195 TATCTGAGAGAGTGCACAAAACAAATTAATTAATCTTACTTAACTGAGAGAG 1254
QY 961 GATTTACGCTGAATATTCACAGATTTGGGAAACATGATGCTTGGCTGCTGCTGCT 1020
DB 1255 GATTTACGCTGAATATTCACAGATTTGGGAAACATGATGCTTGGCTGCTGCTGCT 1314
QY 1021 GTGAAAAGCATCAAAATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1315 GTGAAAAGCATCAAAATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374

RESULT 2
AL512406 80247 bp DNA linear PRI 27-APR-2001
LOCUS
DEFINITION
Human DNA sequence from clone Rpl1-50409 on chromosome 6, complete
sequence.
ACCESSION
AL512406
VERSION
AL512406.14 GI:13897154
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 80247)
REFERENCE
Bates K.
JOURNAL
Direct Submission
Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Apr 30, 2001 this sequence version replaced gi:13396709.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or unsequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30).
An attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone, and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information on the WormPep
database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormPep
This sequence was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
Rpl1-50409 is from the library RCI-11.2 constructed by the group

of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone
RP11-504J9. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP3-381A2 is at 80148 in this sequence.
The true right end of clone RP11-77B15 is at 100 in this sequence.
Location/Qualifiers

FEATURES

```
source
1..80247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-504J9"
/clone_lib="RPC1-11.2"
184..1422
/note="L2 repeat: matches 875..2419 of consensus"
repeat_region
1744..1834
/note="MIR repeat: matches 166..258 of consensus"
repeat_region
2950..2977
/note="14 copies 2 mer ca 100% conserved"
repeat_region
3411..3669
/note="LTP16C repeat: matches 93..364 of consensus"
repeat_region
4372..4704
/note="MIR1A1 repeat: matches 5..365 of consensus"
misc_feature
complement(7821..8406)
/note="match: GSS: Em: A0550748"
8126..8184
/note="MER58B repeat: matches 272..332 of consensus"
8185..8696
/note="MIR2D repeat: matches 1..513 of consensus"
repeat_region
8779..8817
/note="MIR2D repeat: matches 504..542 of consensus"
8819..8974
/note="MER58A repeat: matches 6..162 of consensus"
misc_feature
complement(9443..9891)
/note="match: GSS: Em: B90633"
complement(9452..9883)
/note="match: GSS: Em: A0119695"
9782..9894
/note="MIR repeat: matches 92..205 of consensus"
10089..10212
/note="MIR repeat: matches 135..262 of consensus"
misc_feature
complement(10093..10671)
/note="match: GSS: Em: A0237061"
10223..10282
/note="MIR repeat: matches 203..262 of consensus"
misc_feature
complement(10315..10631)
/note="match: GSS: Em: A0200259"
10805..11186
/note="match: GSS: Em: A0072353"
11657..11844
/note="L2 repeat: matches 2546..2750 of consensus"
misc_feature
complement(11659..12084)
/note="match: GSS: Em: A0518361"
complement(11674..11987)
/note="match: GSS: Em: A0094678"
12083..12588
/note="match: GSS: Em: A0691235"
12596..12810
/note="MIR repeat: matches 23..248 of consensus"
12920..12986
/note="L1MA5 repeat: matches 6228..6295 of consensus"
12987..13169
/note="L1MA5 repeat: matches 5852..6043 of consensus"
13168..13280
/note="L1MA8 repeat: matches 5710..5832 of consensus"
13851..17218
/note="L1P1 repeat: matches 2787..6155 of consensus"
18903..18953
/note="MIR2CB repeat: matches 439..501 of consensus"
18984..19433
/note="MIR2CB repeat: matches 1..460 of consensus"
```

```
misc_feature
complement(19108..19690)
/note="match: GSS: Em: A0563935"
20683..20801
/note="ALUJO/FRAM repeat: matches 169..288 of consensus"
20802..20862
/note="ALU/FLAM repeat: matches 3..66 of consensus"
20912..21033
/note="MIR repeat: matches 114..237 of consensus"
21047..21128
/note="41 copies 2 mer ta 63% conserved"
21676..21859
/note="L2 repeat: matches 2130..2308 of consensus"
21890..23612
/note="L1P repeat: matches 841..2570 of consensus"
23601..26609
/note="L1PA5 repeat: matches 3107..6145 of consensus"
26620..26789
/note="MER5A repeat: matches 1..174 of consensus"
28521..28627
/note="MER39 repeat: matches 5..111 of consensus"
28656..28837
/note="MER39B repeat: matches 54..235 of consensus"
28838..29142
/note="ALUS repeat: matches 4..308 of consensus"
29143..29494
/note="MER39B repeat: matches 235..549 of consensus"
30529..31344
/note="L1PA3 repeat: matches 5331..6146 of consensus"
32234..32340
/note="L2 repeat: matches 2627..2750 of consensus"
32399..32578
/note="MIR repeat: matches 12..203 of consensus"
32816..32979
/note="MIR repeat: matches 91..259 of consensus"
complement(33006..33450)
/note="match: GSS: Em: A0261227"
complement(33058..33444)
/note="match: GSS: Em: A0040453"
33312..33596
/note="L1M4 repeat: matches 3473..3773 of consensus"
34856..34915
/note="30 copies 2 mer aa 66% conserved"
complement(34987..35547)
/note="match: GSS: Em: A0321598
match: STS: Em: G55571"
complement(36488..36936)
/note="match: GSS: Em: B41073"
complement(36654..37090)
/note="match: GSS: Em: A0772584"
37102..37591
/note="match: GSS: Em: A0489567"
complement(37581..38221)
/note="match: GSS: Em: A0418233"
37720..37772
/note="L2 repeat: matches 2657..2710 of consensus"
complement(37778..38215)
/note="match: GSS: Em: A0819773"
38218..38732
/note="match: GSS: Em: A059968"
38861..39164
/note="ALUS repeat: matches 1..305 of consensus"
39612..39653
/note="21 copies 2 mer ac 90% conserved"
40854..40909
/note="L2 repeat: matches 2655..2710 of consensus"
42169..42270
/note="MER5A repeat: matches 1..111 of consensus"
42411..42548
/note="MER5B repeat: matches 21..181 of consensus"
42780..42938
/note="MER5B repeat: matches 17..177 of consensus"
43229..43326
/note="MIR repeat: matches 40..148 of consensus"
```

```
repeat_region      45383. .45428  
/note="ERRB repeat: matches 18. .62 of consensus"  
misc_feature       complement(47284. .47971)  
/note="match: STS: Em:G51995  
match: GSS: Em:AQ269059"  
misc_feature       complement(47537. .48011)  
/note="match: STS: Em:G36899"  
misc_feature       complement(47723. .48014)  
/note="match: STS: Em:G30049"  
misc_feature       47972. .48451  
/note="match: GSS: Em:AQ336527"  
misc_feature       complement(48518. .48985)  
/note="match: GSS: Em:AQ801834"  
repeat_region      48597. .48803
```

Query Match	99.98%	Score 1078.4;	DB 9;	Length 80247;
Best Local Similarity	99.98%	Pred. No. 3e-268;		
Matches 1079; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	1	ATGCATCAACATCCAAAGAAATCTTGCCCATTTTAAATGTCGATTAATCCGGG	60
Db	35559	ATGCATCAACATCCAAAGAAATCTTGCCCATTTTAAATGTCGATTAATCCGGG	35618
OY	61	TGTTTCATGCGATCTGTCATTTACATCAAACTACCAACAGCTGATCTTCACGCA	120
Db	35619	TGTTTCATGCGATCTGTCATTTACATCAAACTACCAACAGCTGATCTTCACGCA	35678
OY	121	ATGGAATCAGCCAGCTCTGTGCTGAAATGAAAACTCTTTCCACCAAACTGATTA	180
Db	35679	ATGGAATCAGCCAGCTCTGTGCTGAAATGAAAACTCTTTCCACCAAACTGATTA	35738
OY	181	TTTATGAAACTACTATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	240
Db	35739	TTTATGAAACTACTATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	35798
OY	241	TCCGCGCAACATGTTCAACATCCAAAGATCCATCTCACAACGAGCGGTACGTAC	300
Db	35799	TCCGCGCAACATGTTCAACATCCAAAGATCCATCTCACAACGAGCGGTACGTAC	35858
OY	301	AACAAATCCGATGCTATTCGATCCATCCAGACATCACTTGTGGATCTGCACAAATTA	360
Db	35859	AACAAATCCGATGCTATTCGATCCATCCAGACATCACTTGTGGATCTGCACAAATTA	35918
OY	361	CTTCAGCAAGCTTAGGCCACCCCTTCCAGAAATGATTTTGATGAATTTGGAATCACA	420
Db	35919	CTTCAGCAAGCTTAGGCCACCCCTTCCAGAAATGATTTTGATGAATTTGGAATCACA	35978
OY	421	CACACTCCCAAAAGAGTGGCATTTGAGCACTTTTAACTGACTCTGACTTACCGCGCT	480
Db	35979	CACACTCCCAAAAGAGTGGCATTTGAGCACTTTTAACTGACTCTGACTTACCGCGCT	36038
OY	481	GATCAGATTTCCAAAGGCTTATGGCTCTTCCGAGGTAACACAAATCCCTTCGATTT	540
Db	36039	GATCAGATTTCCAAAGGCTTATGGCTCTTCCGAGGTAACACAAATCCCTTCGATTT	36098
OY	541	GAAATGCGCAAGCAAGAAATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	600
Db	36099	GAAATGCGCAAGCAAGAAATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	36158
OY	601	GCCAGATCAAGTATTACAAATGAGCTAAGCAAAAGCATTTGAAATCCATACCTAC	660
Db	36159	GCCAGATCAAGTATTACAAATGAGCTAAGCAAAAGCATTTGAAATCCATACCTAC	36218
OY	661	GCAATTTGGAANAATTTCAATGATTAATAATTTGATTCCTACCATATCGCTTGAATTT	720
Db	36219	GCAATTTGGAANAATTTCAATGATTAATAATTTGATTCCTACCATATCGCTTGAATTT	36278
OY	721	TATCTTTCTTTGAAATTCATCCCAAGATTTCAATCAGCGAAAGCGTATACATGCT	780
Db	36279	TATCTTTCTTTGAAATTCATCCCAAGATTTCAATCAGCGAAAGCGTATACATGCT	36338
OY	781	TTTCTGGCTGCTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	840

Db	36339	TTTTGGCTGGCTCTCTACCTGTGTTCTGGGACCACCTGACGAAACATATGGAATTAAT	36398
Qy	841	ATTCCAGCAGATTTCATTTCATGCTGSAAGATTATAACTCTCCAGTAGTACGAAG	900
Db	36399	ATTCCAGCAGATTTCATTTCATGCTGSAAGATTATAACTCTCCAGTAGTACGAAG	36458
Qy	901	TATCTGAGGAACTGACAAAAACAATTAAGTTATACCTGATGTAACCTTAACGTGAGGAAG	960
Db	36459	TATCTGAGGAACTGACAAAAACAATTAAGTTATACCTGATGTAACCTTAACGTGAGGAAG	36518
Qy	961	GATTTCACCTGAATCTTCCAGGATTTTGGAAATCAGATGATGTTGGCTGCATCAT	1020
Db	36519	GATTTCACCTGAATCTTCCAGGATTTTGGAAATCAGATGATGTTGGCTGCATCAT	36578
Qy	1021	GTCGAAAGCATCAAAATATATAGCTCTGCTGTAATTAGAGAAATGCTTTGGAAATTA	1080
Db	36579	GTCGAAAGCATCAAAATATATAGCTCTGCTGTAATTAGAGAAATGCTTTGGAAATTA	36638

RESULT 3	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	SOURCE	
HS0238701	2501 bp	Homo sapiens mRNA for alpha-3-fucosyltransferase.	AJ238701	GI:4741566	alpha-3-fucosyltransferase; FUT9 gene.	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 2501)	Cailliau-Thomaz, A., Collin, P., Candelier, J. J., Balanzino, L., Mennesson, B., Ortol, R. and Mollicone, R.	FUT9 and FUT9 genes are expressed early in human embryogenesis	Glycobiology 10 (8), 789-802 (2000)	2 (bases 1 to 2501)	Mollicone, R.	Direct Submission	Submitted (29-APR-1999)	Mollicone R., Glycobiology, INSERM U504, 16 Av. Paul Vaillant-Couturier, Villejuif, 94807 Cedex, FRANCE	related sequence: AB015426.	Location/Qualifiers	I..2501

```
gene /tissue_type="whole embryo"  
dev_stage="40 to 60 days"  
105..2466  
/gene="FUT9"  
105..1184  
/gene="FUT9"  
/codon_start=1  
/product="alpha-3-fucosyltransferase"  
/protein_id="CAB41890.1"  
/db_xref="GI:4741567"  
/db_xref="SPTREMBL:O9Y231"  
translation="MTSTSKGILRPFLLVCIILIGCFMACLLIYPINSMWIFPMESA  
SYLVKAKNEFFSRKYTFNERTILLVWPFCQETDLSICAMNIIGCHLTVDRIYLKSA  
SHAVLIHHNRISMDLTNLNPQOARPPQPKVMIMNLESPRTPOKSGIEHFLATLYRR  
DSDIQVAGLYLTNTNPVEFEVPSKEKLVCVVSNNNPHAAVKTYNELSGSIIRHY  
GOAFGEVNDKNILPTISACKFYSLSPNSIKDYITEEKLYNPLAGSVYVPLVFGSENN  
YEVIIPADSDIHVEDNSPEELAKYLKEVDKKNNKLYLSTFNMRKDFYNLPFMESHA  
CLACDHVRHQEKSVGNLEKMFVN"  
2192  
/gene="FUT9"  
/evidence=experimental  
2235  
/gene="FUT9"  
/evidence=experimental  
2310  
/gene="FUT9"
```

```

polyA_site      /evidence-experimental
                2466
                /gene="FUT9"
                /evidence-experimental
BASE COUNT      818 a      436 c      422 g      825 t
ORIGIN
Query Match      99.6%; Score 1075.2; DB 9; Length 2501;
Best Local Similarity 99.7%; Pred. No. 2.2e-267;
Matches 1077; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1  ATGACATCAACATCCAAAGAAATCTTCGCCCATTTTAAATGTCGCAATATCCGGGC 60
    |||||||
DB 105 ATGACATCAACATCCAAAGAAATCTTCGCCCATTTTAAATGTCGCAATATCCGGGC 164
QY 61  TGTTCATGGCAGTCCTTCATTTACATCAAACTACCAACAGCTGATCTTCAGTCCA 120
    |||||||
DB 165 TGTTCATGGCAGTCCTTCATTTACATCAAACTACCAACAGCTGATCTTCAGTCCA 224
QY 121 ATGGAATCAGCAGCTCTGCTGTAATAATGAAAAACTCTTTTCCACCAAACTGATAT 180
    |||||||
DB 225 ATGGAATCAGCAGCTCTGCTGTAATAATGAAAAACTCTTTTCCACCAAACTGATAT 284
QY 181 TTTAATGAATACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
    |||||||
DB 285 TTTAATGAATACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
QY 241 TCTGCGCAAGCAATGTTCAACATCCAAAGATGCGCATCTGCACAAAGGACGTTCACTGTAC 300
    |||||||
DB 345 TCTGCGCAAGCAATGTTCAACATCCAAAGATGCGCATCTGCACAAAGGACGTTCACTGTAC 404
QY 301 AACAAATCCCATGCACTTCTGATCATCAACGAGACATGATGGATCTGACAAATTTA 360
    |||||||
DB 405 AACAAATCCCATGCACTTCTGATCATCAACGAGACATGATGGATCTGACTAATTTA 464
QY 361 CCTCAGCAAGCTAGGCCACCTTCCAGAAATGATTTGATGATTTGCAATACCAACT 420
    |||||||
DB 465 CCTCAGCAAGCTAGGCCACCTTCCAGAAATGATTTGATGATTTGCAATACCAACT 524
QY 421 CACACTCCCAAAAGATGGAGCTTGTGAGCACTTGTAACTGACTGACTTACCCCG 480
    |||||||
DB 525 CACACTCCCAAAAGATGGAGCTTGTGAGCACTTGTAACTGACTGACTTACCCCG 544
QY 481 GATTCAGATATCCAAAGTGGCTTATGAGCTTCTGACGCTTAAGCAAAATCCCTTCTGATTT 540
    |||||||
DB 585 GATTCAGATATCCAAAGTGGCTTATGAGCTTCTGACGCTTAAGCAAAATCCCTTCTGATTT 644
QY 541 GAAGTGCACCAAGAAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
    |||||||
DB 645 GAAGTGCACCAAGAAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
QY 601 GCCAGAGTCAAGTATATCAATGAGTACGCAAAAGCAATGCAATACCTACGAGGCAA 660
    |||||||
DB 705 GCCAGAGTCAAGTATATCAATGAGTACGCAAAAGCAATGCAATACCTACGAGGCAA 764
QY 661 GCATTTGGAGAAATATGCTCAATGAAAAATTTGATTTCTTACCAATATCTCTGTAATTT 720
    |||||||
DB 765 GCATTTGGAGAAATATGCTCAATGAAAAATTTGATTTCTTACCAATATCTCTGTAATTT 824
QY 721 TATCTTTCTTTGAAAAATTCATCCCAAGATTTACATACGCAAAAGCTATACATTCCT 780
    |||||||
DB 825 TATCTTTCTTTGAAAAATTCATCCCAAGATTTACATACGCAAAAGCTATACATTCCT 884
QY 781 TTTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
    |||||||
DB 885 TTTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
QY 841 ATTTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 900
    |||||||
DB 945 ATTTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1004
QY 901 TATCTGAAGAAAGTGCACAAAGAAATTAAGTTATACCTTATAGTTTAACTTGAAGGAAG 960
    |||||||

```

```

DB 1005 TATCTGAAGAAAGTGCACAAAGAAATTAAGTTATACCTTATAGTTTAACTTGAAGGAAG 1064
QY 961 GATTCACATGTAATATCTTCACGATTTTGGGAATCAATCATGTTGGCTGCATTCAT 1020
    |||||||
DB 1065 GATTCACATGTAATATCTTCACGATTTTGGGAATCAATCATGTTGGCTGCATTCAT 1124
QY 1021 GTGAAAAAGCATCAAGAAATTAAGTCTGTTGTAATTTAGAGAAATGTTTTGGAATTTA 1080
    |||||||
DB 1125 GTGAAAAAGCATCAAGAAATTAAGTCTGTTGTAATTTAGAGAAATGTTTTGGAATTTA 1184

RESULT 4
AB015426      2139 bp      mRNA      linear      ROD 23-JUN-1999
LOCUS      AB015426
DEFINITION      Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, complete
                cds.
ACCESSION      AB015426
VERSION      AB015426.1 GI:3702718
KEYWORDS      Fut9; alpha1,3-fucosyltransferase IX.
SOURCE      Mus musculus
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                1 (sites)
                Kudo,T., Ikehara,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K.
                and Narimatsu,H.
                Expression cloning and characterization of a novel murine alpha1,
                3-fucosyltransferase, mFuc-TIX, that synthesizes the Lewis x (CD15)
                epitope in brain and kidney
                J. Biol. Chem. 273 (41), 26729-26738 (1998)
                98434588
MEDLINE      2 (bases 1 to 2139)
REFERENCE      Kudo,T. and Narimatsu,H.
AUTHORS      Direct Submission
TITLE      Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science,
                Soka University, Division of Cell Biology; 1-236 Tangi-cho,
                Hachioji, Tokyo 192-8577, Japan (E-mail:tkudo@et.soka.ac.jp,
                Tel:+81-426-91-2495, Fax:+81-426-91-9315)
                Location/Qualifiers
                1..2139
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                115..1194
                /gene="Fut9"
                115..1194
                /gene="Fut9"
                /codon_start=1
                /product="alpha1,3-fucosyltransferase IX"
                /protein_id="BAA33522.1"
                /db_xref="GI:3702719"
                /translation="MTSTSKGILRPFLIYCIILGCFMCLIIYIKPTNSWVSPMESA
                SSVLKMKNFESTKIDYFNETTLVWVMPGQFDTLSCQAMNIOGCHLIDRSIYKX
                SHAVLIHHRDISMDLTNPQCARPPQKIMNLSPLHTPOKSIIEHLFNILTYR
                DSDIDVPEYGLIVSTNPVEFVEPSKEKLVQWVSNWNEHAAVKYVNLSTIEHTY
                GOAFGEVYNDKMLIPTIISTCKYLFSPMSIHRDYITKRLYNALFAGSPVVLGSPREN
                YENYIPADSFIEVEDFNSBELAKYLEKDKNNKLYLSYFNMRKDFYVNLPRFWSHA
                CLACDVKRHHQEKRSVGNLEKWFVN"
                2139
                /note="31 a nucleotides"
polyA_site
BASE COUNT      684 a      395 c      368 g      692 t
ORIGIN
Query Match      86.2%; Score 931.2; DB 10; Length 2139;
Best Local Similarity 91.4%; Pred. No. 3.7e-230;
Matches 987; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1  ATGACATCAACATCCAAAGAAATCTTCGCCCATTTTAAATGTCGCAATATCCGGGC 60
    |||||||
DB 115 ATGACATCAACATCCAAAGAAATCTTCGCCCATTTTAAATGTCGCAATATCCGGGC 174
QY 61  TGTTCATGGCAGTCCTTCATTTACATCAAACTACCAACAGCTGATCTTCAGTCCA 120
    |||||||
DB 175 TGTTCATGGCAGTCCTTCATTTACATCAAACTACCAACAGCTGATCTTCAGTCCA 234

```

[illegible]

KEYWORDS	SOURCE	ORGANISM	REFERENCE	JOURNAL	PUBMED	REFERENCE	TITLE	FEATURES
Norway rat.	Rattus norvegicus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerogamath; Muridae; Murinae; Rattus.	1 (bases 1 to 1128)	Baboval,T., Henion,T., Kinnally,E. and Smith,F.I.			Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX) and comparison of the expression of fuc-rtv and fuc-tix genes during rat postnatal cerebellum development	J Neurosci. Res. 62 (2), 206-215 (2000)
2 (bases 1 to 1128)	Smith,F.I. and Baboval,T.	Direct Submission	Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver Center, 200 Trapelo Rd., Waltham, MA, USA					
1. .1128	/organism="Rattus norvegicus"	/strain="Sprague-Dawley"	/db_xref="taxon:10116"					
49. .1128	/gene="Fut9"							
/gene="Fut9"								
/codon_start=1								
/product="alpha1,3-fucosyltransferase IX"								
/protein_id="AAK16591.1"								
/db_xref="GI:13242184"								
/translation="MTSTSKGILRPPLIVCIILGCPVALILYIKPTMSWFSPESEA SVLTKMFEFRKRDYENETIIYWWPGQFDEITDCOAMNIOGCHITDRLKYK SHAVI.IIHRDI.SMDLTNLPDQAPPPKRYMANSPTLPKSEHLPLATTYRR DSDIQVPGPLVTSPTSPVEPSKSKIKYVSNMNPENHAKRYNLSLSITHTY CGAEFYVNDKNI.IPTISCTCKFLSPENS.IHNDYTERKLYN.FLGSFVVLGSRKRE YNYI.PAASF.IHVEDPNSPSBLAKITILEVDVKNKILTLTFNMKDPVNLDPFWESEA CNAOHVNRHDEIIVSVGLERKMN"								
BASE COUNT	324 a	256 c	228 g	320 t				
ORIGIN								
Query Match	86.1%;	Score 929.6;	DB 10;	Length 1128;				
Best Local Similarity	91.3%;	Pred. No. 9.8e-230;						
Matches 986;	Conservative	0;	Mismatches 94;	Indels 0;	Gaps 0;			
1 ATGACATCAACATCCAAAGAAATCTTGCGCCATTTTAAATGTCTGCAATTACCTGGGC	60							
49 ATGACATCAACATCCAAAGCAATCTTGCGCCATTTTAAATGTCTGCAATTACCTGGGC	108							
61 TGTTCATGCAATGCTTCTTCATTTACATCAAACTACCAACAGCTGGATCTCACTCA	120							
109 TGTTCATGCAATGCTTCTTCATTTACATCAAACTACCAACAGCTGGATCTCACTCA	168							
121 ATGCAATCAGCCAGCTCTGTGCTGAAAGAAATCTTTTTCACACCAAACTGATAT	180							
169 ATGAGATCTGCAAGCTCTGTGCTGAAAGAAATCTTTTTCACACCAAACTGATAT	228							
181 TTTATGAACACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	240							
229 TTTATGAACACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	288							
241 TTTATGAACACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	300							
289 TTTATGAACACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	348							
301 AACCAATCCATGCACTTGTGATCATCAGGAGACATCAGTTGGATGTGACAAATTTA	360							
349 AACCAATCCATGCACTTGTGATCATCAGGAGACATCAGTTGGATGTGACAAATTTA	408							
361 CCTGAGCAAGCTAGGCAACCTTCCAGAAAGGATTTGGATGATTTGGATTCACCACT	420							
409 CCTGAGCAAGCTAGGCAACCTTCCAGAAAGGATTTGGATGATTTGGATTCACCACT	468							

QY 421 CACACTCCCCAAAAGATGGCATTTGAGCACTTTTAACTGACTGACTTACCGCGT 480
 Db 469 CACACCCCCAAAAGATGGCATTTGAGCACTTTTAACTGACTGACTTACCGCGT 528
 QY 481 GATTGAGATATCAAGTGGCTTGTAGCGGTAGGCAACAATCCCTTGTTT 540
 Db 529 GATTGAGATATCAAGTGGCTTGTAGCGGTAGGCAACAATCCCTTGTTT 588
 QY 541 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 600
 Db 589 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 648
 QY 601 GCCAGAGTCAGATATTACATAGCTAGCAAAAAGCATTTGAATTCATACCTAGGCA 660
 Db 649 GCCAGAGTCAGATATTACATAGCTAGCAAAAAGCATTTGAATTCATACCTAGGCA 708
 QY 661 GCATTTGGAGAAATATGTCATATGATAAATTTGATTCCTACCATCTGCTTAAATTT 720
 Db 709 GCATTTGGAGAAATATGTCATATGATAAATTTGATTCCTACCATCTGCTTAAATTT 768
 QY 721 TATCTTCCCTTTGAAAATTCATCAACAAGATTTACATACGGAAGCATATCAATGCT 780
 Db 769 TATCTTCCCTTTGAAAATTCATCAACAAGATTTACATACGGAAGCATATCAATGCT 828
 QY 781 TTTCTGGCTGGCTCTGTACCTGTTGTTCTGGGACATCTAGGAAAACATAGAAATAT 840
 Db 829 TTTTGGCTGGCTCTGTACCTGTTGTTCTGGGACATCTAGGAAAACATAGAAATAT 888
 QY 841 ATTCCAGCATTCATTCATTCATGTAAGATTTAATACCTCCAGTGGAGCAAG 900
 Db 889 ATTCCAGCATTCATTCATTCATGTAAGATTTAATACCTCCAGTGGAGCAAG 948
 QY 901 TATCTGAAGAGTGCACAAAACAATTAATTAATCTAGTACTTAACTGAGGAAG 960
 Db 949 TATCTGAAGAGTGCACAAAACAATTAATTAATCTAGTACTTAACTGAGGAAG 1008
 QY 961 GATTTCAGTGAATATCTTCACAGATTTGGGAATCAACATGATGTTGGCTGCATAT 1020
 Db 1009 GATTTCAGTGAATATCTTCACAGATTTGGGAATCAACATGATGTTGGCTGCATAT 1068
 QY 1021 GTGAAGAAGCATCAAGATATTAAGTCTGTAATTTAGAGAAATGGTTTGAATTTA 1080
 Db 1069 GTGAAGAAGCATCAAGATATTAAGTCTGTAATTTAGAGAAATGGTTTGAATTTA 1128

 RESULT 6
 AB049819 2156 bp mRNA linear ROD 18-JUN-2002
 LOCUS Rattus norvegicus FUT9 mRNA for alpha1,3-fucosyltransferase IX,
 DEFINITION complete cds.
 AB049819
 ACCESSION AB049819.1 GI:13591588
 VERSION
 KEYWORDS Rattus norvegicus cDNA to mRNA.
 SOURCE
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (sites)
 Shimoda, Y., Tajima, Y., Osanai, T., Katsune, A., Kohara, M., Kudo, T.,
 Naitama, S., H., Takashima, N., Ishii, Y., Nakamura, S., Osumi, N. and
 Sanai, Y.
 Pax6 Controls the Expression of Lewis x Epitope in the Embryonic
 Forebrain by Regulating alpha 1,3-Fucosyltransferase IX Expression
 J. Biol. Chem. 277 (3), 2033-2039 (2002)
 JOURNAL 11675393
 PUBMED
 REFERENCE 2 (bases 1 to 2156)
 AUTHORS Sanai, Y.
 TITLE Direct Submission
 JOURNAL Submitted (11-OCT-2000) Yutaka Sanai, Tokyo Metropolitan Institute
 of Medical Science, Department of Biochemical Cell Research,
 Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113-8613, Japan
 (E-mail:sanaie@insoken.or.jp, Tel:81-3-3823-2101(ex.5233),

FEATURES
 source location/Qualifiers
 1..2156
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 gene 346..1425
 /gene="FUT9"
 CDS 346..1425
 /gene="FUT9"
 /codon_start=1
 /product="alpha1,3-fucosyltransferase IX"
 /protein_id="BAB40953.1"
 /db_xref="GI:13591589"
 /translation="MTSTSKGILREPLIYCIIGCEVACLIIYIRPTNSWFSPEMESA
 SSVLKMKNFESRKYDYEFTTLLVWVWPQGFEDLSCAMFIQGHILTDRLSYLNK
 SHAULIHHRDISMDLNLPOOARPPFOKIMWNLSPIHTPOKSGIEHLFNLTLTYR
 DSDIOVPGLFVTSPTPEYFSPKELKAVVSWMNEPHAVKRYNELSIEIHTY
 GQFGEIYNDKULIPTISTCKFLSPENSITHKDYITELKYNFLAGSFPVPLPSREN
 YENYIPDSFIHVEDNPSELAKEYLKEVDKNNKLYLSYFNMRKDFVNLPRFWESHA
 CLACDHVRHOEYKSVGNLEKFWNA"
 BASE COUNT 605 a 470 c 437 g 644 t
 ORIGIN
 Query Match 85.9%; Score 928; DB 10; Length 2156;
 Best Local Similarity 91.2%; Pred. No. 2,5e-229;
 Matches 985; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 1 ATGACATCAACATCCAAAGAAATTTCCGCCATTTTAAATGTCGATATACCTGGGC 60
 Db 346 ATGACATCAACATCCAAAGAAATTTCCGCCATTTTAAATGTCGATATACCTGGGC 405
 QY 61 TGTTCATGCGATGCTTCTCATTTACATCAACAACCTACCAACAGCTGATCTTCACTCA 120
 Db 406 TGTTCATGCGATGCTTCTCATTTACATCAACAACCTACCAACAGCTGATCTTCACTCA 465
 QY 121 ATGGAATCAGCAGCTCTGTGCTGAAAATGAAAACCTCTTTCCACCAAACTGATAT 180
 Db 466 ATGGAATCAGCAGCTCTGTGCTGAAAATGAAAACCTCTTTCCACCAAACTGATAT 525
 QY 181 TTTAATGAACCTCATATCTGCTGGTGGTGGGCAATTTGGGACAGACCTTGACCTTACA 240
 Db 526 TTTAATGAACCTCATATCTGCTGGTGGTGGGCAATTTGGGACAGACCTTGACCTTACA 585
 QY 586 TCTCGCAAGCAATGTTCAACATCCAAAGAGTGCATCTCAACAAGGACCGCTGCTGATAC 645
 Db 241 TCTCGCAAGCAATGTTCAACATCCAAAGAGTGCATCTCAACAAGGACCGCTGCTGATAC 300
 QY 301 AACAAATCCATGCAATTTGATCCATCACCGAGATCAGTTGGATGTGACAAATTTA 360
 Db 646 AACAAATCCATGCAATTTGATCCATCACCGAGATCAGTTGGATGTGACAAATTTA 705
 QY 361 CCTCAGCAAGCTGAGGCGACCTCCAGAAATGATTTGGATGTAATTTGAATCAACCACT 420
 Db 706 CCTCAGCAAGCTGAGGCGACCTCCAGAAATGATTTGGATGTAATTTGAATCAACCACT 765
 QY 766 CACACTCCCCAAAAGATGGCATTTGAGCACTTTTAACTGACTGACTTACCGCGT 480
 Db 825 CACACTCCCCAAAAGATGGCATTTGAGCACTTTTAACTGACTGACTTACCGCGT 885
 QY 885 GATTGAGATATCAAGTGGCTTGTAGCGGTAGGCAACAATCCCTTGTTT 540
 Db 540 GATTGAGATATCAAGTGGCTTGTAGCGGTAGGCAACAATCCCTTGTTT 885
 QY 885 GATTGAGATATCAAGTGGCTTGTAGCGGTAGGCAACAATCCCTTGTTT 885
 Db 885 GATTGAGATATCAAGTGGCTTGTAGCGGTAGGCAACAATCCCTTGTTT 885
 QY 541 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 600
 Db 600 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 945
 QY 945 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 660
 Db 660 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 1005
 QY 1005 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT 720
 Db 720 GAAGTCCCAAGCAAGAAGAAATGTGTGTGGTGTAGTAACTGAAACCTTGACAT

Db 1006 GCATTGGAGAAATATGATGATATAAAATCTGATCCACCATTACTGTAATTT 1065
QY 721 TATCTTCCCTTTGAAATTCATCCAAAGATTCATCAGCAAGAAAGCTATCAATGCT 780
Db 1066 TATCTTCCCTTTGAAATTCATCCAAAGATTCATCAGCAAGAAAGCTATCAATGCT 1125
QY 781 TTTCTGGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1126 TTTTGGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
QY 841 ATTCGACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
Db 1186 ATTCGACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1245
QY 901 TATCTGAGGAAATGTCACAAAACAATTAATTAATTAATTAATTAATTAATTAATTA 960
Db 1246 TATCTGAGGAAATGTCACAAAACAATTAATTAATTAATTAATTAATTAATTAATTA 1305
QY 961 GATTTCACTTAATTCCTACGATTTTGGGATTCATCAGATTCCTTGGCTGAGATCAT 1020
Db 1306 GATTTCACTTAATTCCTACGATTTTGGGATTCATCAGATTCCTTGGCTGAGATCAT 1365
QY 1021 GTGAAAAGGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Db 1366 GTGAAAAGGATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1425

RESULT 7
AF230460 1705 bp mRNA linear ROD 12-JUL-2000
LOCUS DEFINITION Crictellus griseus alpha(1.3)fucosyltransferase type IX (Fut9)
ACCESSION AF230460
VERSION AF230460.1 GI:9049663
KEYWORDS Chinese hamster.
SOURCE Crictellus griseus
ORGANISM Crictellus griseus
REFERENCE 1 (bases 1 to 1705)
AUTHORS Patnaik S.K., Zhang A., Shi S. and Stanley P.
TITLE Chinese hamster ovary glycosyltransferases expressed by the gain-of-function
JOURNAL Arch. Biochem. Biophys. 375 (2), 322-332 (2000)
MEDLINE 2018935
PUBMED 10700388
REFERENCE 2 (bases 1 to 1705)
AUTHORS Patnaik S.K., Shi S. and Stanley P.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of
Medicine, Chanin 516, 1300 Morris Park Avenue, New York, NY 10461,
USA
FEATURES
source Location/Qualifiers
1..1705
/organism="Crictellus griseus"
/db_xref="taxon:10029"
1..1705
/gene="Fut9"
/note="cgbf9"
62..1141
/gene="Fut9"
/note="Fuc-TIX"
/codon_start=1
/product="alpha(1.3)fucosyltransferase type IX"
/protein_id="AAF82412.1"
/db_xref="GI:9049664"
/translation="MTSTSKGILRPLIVCVIACMACLLIYIKPTNWSVPSMESA
SVLTKMKNEFTKDYENETILVWVPFGOTDLSCAMFIQCCHLTTRSLYRK
SHAVLIHRDLSMDLTNLPQARPPPKWIMMLSEPTHPKSGIEHLFTLYRK
DSDIQVPGFLVSNPVEFVPEPKELVQVSNMNPBARVKKYNEKSEIETHY
GOAFGEYENKNIIPITJSTCKEYSPENSIRHKYITEXLIYNAFLAGSVYVGSREN
YENITPADSFIVHEDYNSPSELAKYLEVDKNNKILISTFNNRKDPTVNLPRFMESHA

BASE COUNT 515 a 336 c 320 g 534 t
ORIGIN
Query Match 84.9% Score 916.8; DB 10; Length 1705;
Best Local Similarity 90.6%; Pred. No. 2e-226;
Matches 978; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 1 ATGACATCAACATCCAAAGAAATTTCTGCCCATTCTTAATGTCTGCAATTAATCTGGCC 60
Db 62 ATGACATCAACATCCAAAGAAATTTCTGCCCATTCTTAATGTCTGCAATTAATCTGGCC 121
QY 61 TGTTCATGCGCATGCTTCTCATTTTACATCAACACCTACCAACACCTGATCTTCACTGCA 120
Db 122 TGTTCATGCGCATGCTTCTCATTTTACATCAACACCTACCAACACCTGATCTTCACTGCA 181
QY 121 ATGGAATCAGCCAGCTCTGTGCTGAAAATGAAAACCTTTTCCACCAAACTGATTAAT 180
Db 182 ATGGAATCAGCCAGCTCTGTGCTGAAAATGAAAACCTTTTCCACCAAACTGATTAAT 241
QY 181 TTTAATGAACATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 242 TTTAATGAACATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
QY 241 TCTGTCACGCAATGTTCAACATCCAGAGATCCATCTCACACGACCGTTCACGTGAC 300
Db 302 TCTGTCACGCAATGTTCAACATCCAGAGATCCATCTCACACGACCGTTCACGTGAC 361
QY 301 AACAAATCCCATGCAAGTTCTGATCCATCCAGAGATCCATCTCACACGATCTGACAAATTTA 360
Db 362 AACAAATCCCATGCAAGTTCTGATCCATCCAGAGATCCATCTCACACGATCTGACAAATTTA 421
QY 361 CCTCAGCAAGCTAGCCACCTCTCCAGAAATGATTTGGATGAATTTGAATCACCACACT 420
Db 422 CCTCAGCAAGCTAGCCACCTCTCCAGAAATGATTTGGATGAATTTGAATCACCACACT 481
QY 421 CACACATCCCAAAAGAGTGGCATTTGAGCACTTTGTAACCTGATCTGATTCAGCCGCT 480
Db 482 CACACATCCCAAAAGAGTGGCATTTGAGCACTTTGTAACCTGATCTGATTCAGCCGCT 541
QY 481 GATTCAGATATCCAAATGCTTATGCTTATGCTGATGAGCAACCAATTCCTTCGTGTTT 540
Db 542 GATTCAGATATCCAAATGCTTATGCTTATGCTGATGAGCAACCAATTCCTTCGTGTTT 601
QY 541 GAAGTGCACAGCAAGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 602 GAAGTGCACAGCAAGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 601 GCCAGATCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 662 GCCAGATCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
QY 661 GCATTTGGCAATATGCTCAATGATTAATTTGATTTGCTTCCACATTCCTGCTGTAATTT 720
Db 722 GCATTTGGCAATATGCTCAATGATTAATTTGATTTGCTTCCACATTCCTGCTGTAATTT 781
QY 721 TATCTTCCCTTTGAAATTCATTCATCCAAAGATTCATCAGCAAGAAAGCTATCAATGCT 780
Db 782 TATCTTCCCTTTGAAATTCATTCATCCAAAGATTCATCAGCAAGAAAGCTATCAATGCT 841
QY 781 TTTCTGGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 842 TTTCTGGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
QY 841 ATTCGACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
Db 902 ATTCGACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 961
QY 901 TATCTGAGGAAATGTCACAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
Db 962 TATCTGAGGAAATGTCACAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1021
QY 961 GATTTCACTTAATTCCTACGATTTTGGGATTCATCAGATTCCTTGGCTGAGATCAT 1020

|||||
Db 1022 GATTCTACTGTAACCTCCACGATTGCGATCATGATGCTGCTGTGATCAC 1081
Oy 1021 GTAAAGGATCAAGATATAGTCTGTGTAATTTAGAGAAATGCTTTGGAATTA 1080
Db 1082 GTAAAAAGCACCAGAGATAGTCTGTGGTAATTTAGAGAAATGCTTTGGAATTA 1141

RESULT 8
AB035906 1707 bp DNA linear VRT 05-JAN-2002
LOCUS Gallus gallus gene for CFUT9, complete cds.
ACCESSION AB035906
VERSION AB035906.1 GI:18146865
KEYWORDS CFUT9.
SOURCE Gallus gallus brain DNA.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE
AUTHORS Kaneko,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
TITLE 1 (sites)
JOURNAL The evolutionary history of glycosyltransferase genes
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 1707)
TITLE Direct Submission
JOURNAL Kaneko,M., Saitou,N. and Kitano,T.
Submitted (17-DEC-1999) Mlka Kaneko, National Institute of Genetics, Laboratory of Evolutionary Genetics; Yata 1111, Mishima, Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp, Tel:81-539-81-6790, Fax:81-539-81-6789)
Sequence updated (29-Feb-2000).
COMMENT Location/Qualifiers
FEATURES
source 1. 1707
/organism="Gallus gallus"
/db_xref="taxon:9031"
/cisse_type="brain"
268. 1347
/gene="cfuc-TIX"
268. 1347
/gene="cfuc-TIX"
268. 1347
/gene="cfuc-TIX"
/odon_start=1
/product="CFUT9"
/protein_id="BAB82489.1"
/db_xref="GI:18146866"
/translation="MTSRSGIFRPFIVLIGCFMAFILIKRPSWISNPESA
SSVLKMFSSKSDNENETILVIMPRGQPDLSQCFNIGHCLTIDRSYLNK
SHAVLIHRDINMDLNLPOARPPFORIMWNLSPHTPOKSGTEHLNLTIRK
DSDIQVPGSMIVGATFYEPKSNLVWVSWNEHARVXKYNELSYIEHTY
QAVFGDYNDKHLPTLSSCKFLSFENSIHKDYTERLYNALAGSVPLVGSREN
YENYIPADSFHVEDFLSPRELALEYLMLDKNNKMYLSYFWMRKDFSVHLPKFMSHA
CLACDHVRHDEYSIGNLEKMFNN"
BASE COUNT 517 a 338 c 329 g 521 t 2 others
ORIGIN

Query Match 72.6%; Score 783.6; DB 5; Length 1707;
Best Local Similarity 82.8%; Pred. No. 5; 5e-192;
Matches 894; Conservative 1; Mismatches 185; Indels 0; Gaps 0;

Oy 1 ATGACATCAACATCCAAAGGAATTCGCCCATTTTATGTCATTAATCTGGGC 60
Db 268 ATGACATCAACATCCAAAGGAATTCGCCCATTTTATGTCATTAATCTGGGC 327
Oy 61 TCTTCATGCGATGCTCTGCTGTAAGATGAAGAACTTCTTCCACCAAACTGATAT 180
Db 328 TCTTCATGCGATGCTCTGCTGTAAGATGAAGAACTTCTTCCACCAAACTGATAT 387
Oy 121 ATGGAATCAGCAGCTCTGCTGTAAGATGAAGAACTTCTTCCACCAAACTGATAT 180
Db 388 GTGGAATCAGCAGCTCTGCTGTAAGATGAAGAACTTCTTCCACCAAACTGATAT 447
Oy 181 TTTAATGAACACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
|||||

Db 448 TTTAATGAACACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
Oy 241 TCTGCGAAGCAATGTTCAACATCCAAAGGATGCCATTCACACAGGACGCTGATAC 300
Db 508 TCTGCGAAGCAATGTTCAACATCCAAAGGATGCCATTCACACAGGACGCTGATAC 567
Oy 301 AACAAATCCATGACGTTCTGATCCATCAGCAGACATCAGTGGATGCAAAATTA 360
Db 568 AACAAATCCATGACGTTCTGATCCATCAGCAGACATCAGTGGATGCAAAATTA 627
Oy 361 CCTGACCAAGCTAGGCGCCACCTTCCAGAAATGATTTGATGTAATTTGAAATCACCAC 420
Db 628 CCTGACCAAGCTAGGCGCCACCTTCCAGAAATGATTTGATGTAATTTGAAATCACCAC 687
Oy 421 CACACTCCCAAAAGATGAGCATGAGCAGCTGTTAACTGACTGCTGACTACCGCGCT 480
Db 688 CATACTCCCAAAAGATGAGCATGAGCAGCTGTTAACTGACTGCTGACTACCGCGCT 747
Oy 481 GATTCAGATATCCAAAGTGGCTTATGCTTCTTACAGGTAAAGCACAATCCCTGCTGTT 540
Db 748 GATTCAGATATCCAAAGTGGCTTATGCTTCTTACAGGTAAAGCACAATCCCTGCTGTT 807
Oy 541 GAAGTCCCAAGCAAGAAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 808 GAAGTCCCAAGCAAGAAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
Oy 601 GCCAGAGCAATGATTTACATGAGTACGCAAGCAAGCAATGATTTACATGAGTACGCA 660
Db 868 GCTGCAAGTAAATATTTACATGAGTACGCAAGCAAGCAATGATTTACATGAGTACGCA 927
Oy 661 GCATTTGAGAAATATGCTATGATTAATTTGATTCCTACATGCTGCTGCTGCTGCTGCT 720
Db 928 GCTTTGAGAAATATGCTATGATTAATTTGATTCCTACATGCTGCTGCTGCTGCTGCT 987
Oy 721 TATCTTCTCTTGAATATTCATCCACAGATTTACATGCTGCTGCTGCTGCTGCTGCT 780
Db 988 TATCTTCTCTTGAATATTCATCCACAGATTTACATGCTGCTGCTGCTGCTGCTGCT 1047
Oy 781 TTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1048 CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107
Oy 841 ATTCAGCAGATTCATTCATGATGGAATTAATCTCCAGAGAGCTGAGTACGAAG 900
Db 1108 ATTCAGCAGATTCATTCATGATGGAATTAATCTCCAGAGAGCTGAGTACGAAG 1167
Oy 901 TATTCGAGGAAGTGCACAAAATTAATGATTAATCTTACCTTACCTTACCTGAGGAAG 960
Db 1168 TATTCGAGGAAGTGCACAAAATTAATGATTAATCTTACCTTACCTTACCTGAGGAAG 1227
Oy 961 GATTCACAGTAAATCTCCAGATTTGGGAATCAGATGCTGCTGCTGCTGCTGCTGCT 1020
Db 1228 GATTCACAGTAAATCTCCAGATTTGGGAATCAGATGCTGCTGCTGCTGCTGCTGCT 1287

Oy 1021 GTAAAGGATCAAGATATAGTCTGTGTAATTTAGAGAAATGCTTTGGAATTA 1080
Db 1288 GTAAAGGATCAAGATATAGTCTGTGTAATTTAGAGAAATGCTTTGGAATTA 1347
|||||

RESULT 9
AB035905 1080 bp DNA linear VRT 05-JAN-2002
LOCUS Xenopus laevis gene for xFUT9, complete cds.
DEFINITION Xenopus laevis
ACCESSION AB035905
VERSION AB035905.1 GI:18146863
KEYWORDS xFUT9.
SOURCE Xenopus laevis DNA.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.
REFERENCE
AUTHORS Kaneko,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
1 (sites)

TITLE	The evolutionary history of glycosyltransferase genes
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1080)
AUTHORS	Kaneko,M., Saitou,N. and Kitano,T.
TITLE	Direct Submission
JOURNAL	Submitted (17-DEC-1999) Mika Kaneko, National Institute of

Genetics, Laboratory of Evolutionary Genetics; Yata 1111, Mishima, Shizuoka 411-8540, Japan (E-mail: mkaneko@med.id.yamagata-u.ac.jp), Tel: 81-559-81-6790, Fax: 81-559-81-6789)
Sequence updated (29-Feb-2000).

FEATURES	Location/Qualifiers
source	1. .1080

```

/organism="Xenopus laevis"
/db_xref="taxon:835"
1..1080
/gene="xFuc-TIX"
CDS
1..1080

```

```

/gene="xPuc-T1x"
/codon_start=1
/product="xPucT9"
/protein_id="BAB82488.1"
/db_xref="GI:18146864"
/translation="MTSASKGIILRAFLAVSGVILICPMILCITYKPTNNMISPTIES
HSLTKKRNFESSKONYNENETVYLIVMVEFGOTPELKSOTPEFNIGCHLETTDALYAL
SHVLLKHHRDIDMDLINLPOLRPFQKIMMNNSEPHHPKSGIEHLNLTLYR
DSIOIYGVGFMYVSTKRPDFEYVPSKDRKCVVCWVSNMDEHARVYVVELIKYIEIT
GQAFGEYSDKSLPTPTIASKCFYSEFNSIKHVDITKELYALLAGSVPIVLDPPR
YEVYIPADSFIEVEDFLSPRELADHLMLLNKDTBQLYLYVMKRKHFLVNMDEHESHS
CIACDVKRHKQYKSVGNLEKFMVN"

```

Query Match	62.7%	Score 676.8	DB 5	Length 1080
Best Local Similarity	76.7%	Pred. No. 2.3e-164		
Matches 828: Conservative	0	Mismatches 252	Indels 0	Gaps 0

[illegible][illegible]

121	ATGGCAATCAGCCACGCTCTGTGCTGATAAAATGAAAAACTTCTTTCCACCAAAACTGATTAT	180
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		
201		
202		
203		
204		
205		
206		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		
239		
240		
241		
242		
243		
244		
245		
246		
247		

DB	121	ATCGAATCAGCAGATTCATGATCCCTTAAAAAGAAATTTTTCCTCCCTAAACAGAAATTAT	180
DY	181	TTTAATGAACATCATTCCTGTGTGGGTGTGGCCATTGTGGCAGACCTTGACCTTACA	240
DB	181	TTTAATGAACATCATTCCTGTGTGGGTGTGGCCATTGTGGCAGACCTTGACCTTACA	240

DB	181	YACAAAGAAACCGTCGTCCTAAATTTGGGCTGTGTGGCCCTTTGGGACAGCTTTGAACGTGAAA	240
OY	241	TCCTGCCAAGCAATGTTCAACATCCAAAGGATGTCCTCTCAACAGGACCGTTCACTGTAC	300

DB 241 TCGTGGCCAGACATGTTTATATTTATAGGSGTGCCATCTTACTACGACCGGGCCTCTAC 300

QY 301 AACCAATCCCATGAGTCTGATCCATCACCAGACATCAGTTGGGATCTGACAAATTTA 360

Db 301 AACAGTCTCATGCTGTCTCTCTTCACCCACAGACATTTAGCTGGGATCTATTAACTG 360

Qy 361 CCTCAGCAAGCTAGGCCACCCCTTCAGAAATGGATTTGGATGAATTTGGAAATCACCACCT 420

Dd	361	CCCCCAACGCTGAGGGCCACCTTTTCCAGAAATGCAATTTGCAATGAAATCTGGAATCACCACACA	420
Dy	421	CACACTCCCCAAAAGAGTGGCATTTGAGCAGCTGTTTTAACTGTGACTGTGACTTACCGCGGT	480

Db	421	CATACGCCACAAAAAGTGGCATTTAGGACACTATTTAACCTGACCTGACTTACAGGCGT	480
Qy	481	GATTACAGATATCCAAAGTGCCTTATGGCTTCTTGACGGTAAGCACAAATCCCTTCGTGTTT	540

Db	481	GACTCAGATATCCAAAGTGCCCTATGGCTTCATGTCGTAAAGCACAAACCAATTTGACTTT	540
QY	541	GAAATGCCAAGCAAGAAGAAATGGTGTGCTGGGTTGTGAGTAAGTGAACCCGTGACAT	600

DB	Sequence	Position
Db	541 GAGGTGCCAAGCAAGACAGTTGCTGTTGGGTTGTAGCAACTGSAACCCAGAACAT	600
QY	601 GCCAGAGTCACAGTATTACATGAGCTTAAGCAAAAGCATTTGAATCCATACCTACGGGCAA	660

Db	601	GCACAGACCAATATTCAATGAACTATCAATACATTGAATTAATACCTATATGGCAA	660
Qy	661	GCATTTGGAGATATGCAATGATAAATAATTGATTCCTACCATATGCTGTAAATTT	720

Db 661 GCCTTTCGTGAATACCGCAGCGCATTAAGAAGCCCTGCTTCCACATATTGCATCTTCCAAATTT 720

Qy 721 TATCTTCCCTTGAAGAAATTCATCCACAGAAGATTACATCAGGAAAAGCTATACAAATGCT 780

D_b 721 TATCTGTCITTTGGAAAATTCCATTCACAAGATTATATCAGAAAAACTTATAATGCA 780
 |||||
D_y 781 TTCTCGCTGGCTCTGTACCTGTGTTCTGGACCACCTAGGGAAACCTATGCAGAA'TTAT' 840

Db 781 TTAATGGCTGGGTCAGTGCACATTTGTAATTGGGACCTCTAGAGAAAACTATGAGAACTAC 840

Db 841 ATTCCTGAGATTCCTTATCCACGTCGGAAGATTTCCTCTCCAAAGAAATTAGCTCAC 900

QY 901 TATCTGAAGGAAGTCGCAAAAAACAATAGTTATACCTTAGTTACTTTAACTGAGAGAG 960

Db 901 CATCTTAAATGCTGACCAAGATACAGAACAGTACTGAGCTACTACACACTGGAGGAA 960
 961 GATTTCAGCTGTAAATCTTCCACGATTTTGGGAATCACATGATGTTGGCTTGGGATCAT 1020

Accession	Sequence	Length
961	CAATTTACGTCATATAGCCCTCACTTCTGGGAATCCATGCATGCTTAAGCGTGTGACCAT	1020
1021	GTGAAAAGGCATCAAGAGATATATAGTCTGTTGGTAAATTTAGAGAAATGCTTTGGCAATTA	1080

Db 1021 GTGAAGGCATCAAGATACAAATCATCGGCAATCTAGAGAAAGTGGTTCTGGAACTAA 1080

RESULT 10	AC100674	LOCUS	AC100674	71396 bp	DNA	linear	HTG 22-NOV-2001
DECONTAMINATION	AC100674	AC100674	AC100674	71396 bp	DNA	linear	HTG 22-NOV-2001

```

DEFINITION
MUS MUSCULUS CLONE R423-16/K42, LOW-PASS SEQUENCE SAMPLING.
ACCESSION
AC100674
VERSION
AC100674.1 GI:17048040
KEYWORDS
HTG; HTGS_PHASEO.

```

SOURCE	ORGANISM
house mouse.	<i>Mus musculus</i>
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 71396)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Mus musculus, clone RP23-167K24
JOURNAL Unpublished

REFERENCE
AUTHORS
2 (bases 1 to 71396)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, U., Gardyna, S., Ginde, S., Gord, S., Govette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Hearford, A., Horton, J., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K., Llamas, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Macdonald, C., MacDonald, P., Mator, T., Marquis, N., Matthews, C.

McCathay, M., McEwan, P., McKernan, K., McSheeters, R., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T., Portnoo, K., Rumburg, D., Siano, N., Bellamy, R.

Oliver, O., Petersson, A., Fumkhan, F., Flettle, N., Pollard, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,

Strauss, N., Sudirmanian, A., Ialamas, J., Vesiaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,

TITLE
JOURNAL
COMMENT

zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L15932
Center Clone name: 167_K_24

* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 788: contig of 788 bp in length
789 888: gap of 100 bp
889 1702: contig of 814 bp in length
1703 1802: gap of 100 bp
1803 2642: contig of 840 bp in length
2643 2742: gap of 100 bp
2743 3637: contig of 895 bp in length
3638 3737: gap of 100 bp
3738 4635: contig of 898 bp in length
4636 4735: gap of 100 bp
4736 5657: contig of 922 bp in length
5658 5757: gap of 100 bp
5758 6692: contig of 935 bp in length
6693 6792: gap of 100 bp
6793 7703: contig of 911 bp in length
7704 7803: gap of 100 bp
7804 8742: contig of 939 bp in length
8743 8842: gap of 100 bp
8843 9787: contig of 945 bp in length
9788 9887: gap of 100 bp
9888 10819: contig of 932 bp in length
10820 10919: gap of 100 bp
10920 11808: contig of 889 bp in length
11809 11908: gap of 100 bp
11909 12830: contig of 922 bp in length
12831 12930: gap of 100 bp
12932 13847: contig of 917 bp in length
13848 13947: gap of 100 bp
13948 14886: contig of 939 bp in length
14887 14986: gap of 100 bp
14987 15911: contig of 925 bp in length
15912 16011: gap of 100 bp
16012 16903: contig of 882 bp in length
16904 17003: gap of 100 bp
17004 17906: contig of 903 bp in length
17907 18006: gap of 100 bp
18007 18931: contig of 925 bp in length
18932 19031: gap of 100 bp
19032 19945: contig of 914 bp in length
19946 20045: gap of 100 bp
20046 20951: contig of 906 bp in length
20952 21051: gap of 100 bp
21052 21949: contig of 898 bp in length
21950 22049: gap of 100 bp
22050 22950: contig of 901 bp in length
22951 23050: gap of 100 bp

23051 23945: contig of 895 bp in length
23946 24045: gap of 100 bp
24046 24955: contig of 910 bp in length
24956 25055: gap of 100 bp
25056 25980: contig of 925 bp in length
25981 26080: gap of 100 bp
26081 27052: contig of 972 bp in length
27053 27152: gap of 100 bp
27153 28067: contig of 915 bp in length
28068 28167: gap of 100 bp
28168 29076: contig of 909 bp in length
29077 29176: gap of 100 bp
29177 30093: contig of 917 bp in length
30094 30193: gap of 100 bp
30194 31099: contig of 906 bp in length
31100 31199: gap of 100 bp
31200 32097: contig of 898 bp in length
32098 32197: gap of 100 bp
32198 33113: contig of 916 bp in length
33114 33213: gap of 100 bp
33214 34125: contig of 912 bp in length
34126 34225: gap of 100 bp
34226 35133: contig of 908 bp in length
35134 35233: gap of 100 bp
35234 36150: contig of 917 bp in length
36151 36250: gap of 100 bp
36251 37171: contig of 921 bp in length
37172 37271: gap of 100 bp
37272 38159: contig of 888 bp in length
38160 38259: gap of 100 bp
38260 39156: contig of 897 bp in length
39157 39256: gap of 100 bp
39257 40188: contig of 932 bp in length
40189 40288: gap of 100 bp
40289 41210: contig of 922 bp in length
41211 41310: gap of 100 bp
41311 42222: contig of 912 bp in length
42223 42322: gap of 100 bp
42323 43218: contig of 896 bp in length
43219 43318: gap of 100 bp
43319 44233: contig of 915 bp in length
44234 44333: gap of 100 bp
44334 45246: contig of 913 bp in length
45247 45346: gap of 100 bp
45347 46261: contig of 915 bp in length
46262 46361: gap of 100 bp
46362 47279: contig of 918 bp in length
47280 47379: gap of 100 bp
47380 48276: contig of 897 bp in length
48277 48376: gap of 100 bp
48377 49282: contig of 906 bp in length
49283 49382: gap of 100 bp
49383 50305: contig of 923 bp in length
50306 50405: gap of 100 bp
50406 51314: contig of 909 bp in length
51315 51414: gap of 100 bp
51415 52335: contig of 921 bp in length
52336 52435: gap of 100 bp
52436 53163: contig of 728 bp in length
53164 53263: gap of 100 bp
53264 54210: contig of 947 bp in length
54211 54310: gap of 100 bp
54311 55199: contig of 889 bp in length
55200 55299: gap of 100 bp
55300 56212: contig of 913 bp in length
56213 56312: gap of 100 bp
56313 57287: contig of 975 bp in length
57288 57387: gap of 100 bp
57388 58245: contig of 858 bp in length
58246 58345: gap of 100 bp
58346 59236: contig of 891 bp in length
59237 59336: gap of 100 bp
59337 60272: contig of 936 bp in length

```

* 60373 60372: gap of 100 bp
* 60373 61289: contig of 917 bp in length
* 61290 61389: gap of 100 bp
* 61390 62299: contig of 910 bp in length
* 62300 62399: gap of 100 bp
* 62400 63322: contig of 923 bp in length
* 63323 63422: gap of 100 bp
* 63423 64336: contig of 914 bp in length
* 64337 64436: gap of 100 bp
* 64437 65354: contig of 918 bp in length
* 65355 65454: gap of 100 bp
* 65455 66352: contig of 898 bp in length
* 66353 66452: gap of 100 bp
* 66453 67362: contig of 910 bp in length
* 67363 67462: gap of 100 bp
* 67463 68370: contig of 908 bp in length
* 68371 68470: gap of 100 bp
* 68471 69379: contig of 909 bp in length
* 69380 69479: gap of 100 bp

```

Query Match 59.2%; Score 639; DB 2; Length 71396;
 Best Local Similarity 85.0%; Pred. No. 1,2e-154;
 Matches 764; Conservative 0; Mismatches 122; Indels 13; Gaps 4;

```

QY 72 ATGTCCTCATTTTACATCAACCAAGCTGATCTTCACTTCAATGCAATGCAAC 131
DB 65456 ATCTCTGCCCTGTGGTGGATCCCAACAGCTGGCTTCACTTCAATGCAATGCAAC 131
QY 132 CAGCTGTGCTGAAAAAAGTCTTCTTCCACCAAACTGATTTATTTAATGAAC 191
DB 65516 AAGTTCTGCTCTAAAAATGAAAAATTTCTTCCACAAAACTGATTTATTTAATGAAC 191
QY 192 TACTATCTGCTGTGGTGGTGGATTTGGGAGACCTTGCATTAATCTGCGCAAGC 251
DB 65576 TACATCTGCTGTGGTGGTGGATTTGGGAGACCTTGCATTAATCTGCGCAAGC 251
QY 252 AATGTTCAATCAAGGATGCCATCTGCACAGGACCGTTCACTTCAATCAATCAAG 311
DB 65636 AATGTTCAATCAAGGATGCCATCTGCACAGGACCGTTCACTTCAATCAATCAAG 311
QY 312 TGCAGTCTGATTCATCAGCAGACATGAGTGGATCTGACAAATTTAATCTGCAAGC 371
DB 65696 TGCAGTCTGATTCATCAGCAGACATGAGTGGATCTGACAAATTTAATCTGCAAGC 371
QY 372 TAGGCACTCTTCCAGAAATGATTTGATGATTTGGAATCACCACACTGACCTCCA 431
DB 65756 CAGGCACTCTTCCAGAAATGATTTGATGATTTGGAATCACCACACTGACCTCCA 431
QY 432 AAAGAGTGGCATTTGACCACTTGTAACTGACTGACTTACCGCGCTGATTCAGATAT 491
DB 65816 AAAGAGTGGCATTTGACCACTTGTAACTGACTGACTTACCGCGCTGATTCAGATAT 491
QY 492 CCAAGTCTTATGCTTCTTCAAGGTAAGCAAAATCCCTGCTTGAAGTGCACAG 551
DB 65876 CCAAGTCTTATGCTTCTTCAAGGTAAGCAAAATCCCTGCTTGAAGTGCACAG 551
QY 552 CAAAGAGAAATTTGCTGCTGGTGTGATGATGAGTGAACCGCTGAGCATGCAAGTCAA 611
DB 65936 CAAAGAGAAATTTGCTGCTGGTGTGATGATGAGTGAACCGCTGAGCATGCAAGTCAA 611
QY 612 GTATTTCATTAAGCTTAAGCAAAAGCAATTAATCAATCAATCAATCAATCAATCA 671
DB 65996 GTATTTCATTAAGCTTAAGCAAAAGCAATTAATCAATCAATCAATCAATCAATCA 671
QY 672 ATATGTCATGATTAATAATTTGATTCCTATCTGCTGATGATTTAATTTATCTTCT 731
DB 66056 ATATGTCATGATTAATAATTTGATTCCTATCTGCTGATGATTTAATTTATCTTCT 731
QY 732 TGAATATTCATCCACAGGATTTACATCAGGAAAGATATACATGCTTTCTGCTG 791
DB 66116 TGAATATTCATCCACAGGATTTACATCAGGAAAGATATACATGCTTTCTGCTG 791
QY 792 CTCGTACACGTGTTCTGGACCATCTGAGGAAAGAACTATAGAAATTTATTCACAGCA 851

```

```

DB 66167 TTCATCTACCTTGTCCCGGGTCCATTCAGGAAACTATGAAATTAATTCACAGCTGA 66226
QY 852 TTCATCTACCTTGTCCCGGGTCCATTCAGGAAACTATGAAATTAATTCACAGCTGA 908
DB 66227 TTCATCTACCTTGTCCCGGGTCCATTCAGGAAACTATGAAATTAATTCACAGCTGA 908
QY 909 GGAATTCACCAAAACATAGATTTATCTTGAATTAATTCAGGAAAGATTTAC 966
DB 66287 GGAATTCACCAAAACATAGATTTATCTTGAATTAATTCAGGAAAGATTTAC 966

```

RESULT 11
 AB023627
 LOCUS
 DEFINITION
 Danto rerio zrf1 gene for alpha(1,3)fucosyltransferase, complete
 ACCESSION
 AB023627
 VERSION
 AB023627.1 GI:4587295
 KEYWORDS
 alpha(1,3)fucosyltransferase.
 SOURCE
 Danto rerio DNA.
 ORGANISM
 Danto rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danto.
 REFERENCE
 1 (sites)
 Kageyama, N., Natsuka, S. and Hase, S.
 Molecular cloning and characterization of two zebrafish
 alpha(1,3)fucosyltransferase genes developmentally regulated in
 embryogenesis
 J. Biochem. (1999) In press
 2 (bases 1 to 1567)
 Kageyama, N., Natsuka, S. and Hase, S.
 Direct Submission
 Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate
 School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034,
 Japan (E-mail:natsuka@chem.sci.osaka-u.ac.jp, Tel:81-6-6850-5381,
 Fax:81-6-6850-5383)
 FEATURES
 Location/Qualifiers
 source
 1..1567
 /organism="Danto rerio"
 /db_xref="taxon:7955"
 177..1352
 /gene="zrf1"
 177..1352
 /gene="zrf1"
 /codon_start=1
 /product="alpha(1,3)fucosyltransferase"
 /protein_id="BA076706.1"
 /db_xref="GI:4587296"
 translation="MDKILTPPSNAOKVIAIPMLISFYCIFVYVYNNPTFFKFP
 AVDHNTSCSTEDLDLVKRYKCTTKRASNPSTRAAPAPATKDDPFIIL
 MYVFGARFELDVGSGENIHGCHLIDRSMIOKHGMFMRHRLSDLPAPAD
 KWIIMNPSNSHNSRSLINDPFLITSPRSKIDRIVYGLIDALDEOKNPTPKD
 KLVCMIVSNDEHNKRISQYVNLVYHIVLEAYGSDIPYNNRVDDIDYGVNCSRPYLF
 ENSHIDYPERKIFNDLALCTVYLIGSRDNYPERITRDAPFIVDDYDPPPKELADL
 KSDQNDQVROKTFNPKRIFISMSSPELEHRCRCDYLKNNKHYIAIKDKGFWG"
 BASE COUNT
 499 a 334 c 293 g 441 t
 ORIGIN

Query Match 25.7%; Score 277.8; DB 5; Length 1567;
 Best Local Similarity 58.5%; Pred. No. 3.3e-61;
 Matches 527; Conservative 0; Mismatches 362; Indels 12; Gaps 2;

```

QY 183 TAATGAATCTATCTGCTGTGGTGGTGGATTTGGGAGACCTTGAATCTTACATC 242
DB 461 TCAGAAACATCTCTGATCTGCTGATCTGCTGATCTGCTGATCTGCTGATCTGCTG 520
QY 243 CTGCAACAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 302
DB 521 GTGTGTTGGAGTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 580
QY 303 CAATCCATGACATGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 362

```

Db 581 GAAAGCCGATGAGTTTATGTTTACACAGAGATCTGAGCGG-----ATTGGC 631
 Qy 363 TCAGCAAGTACAGCCCTTCACAAATGATTTGATGTAATTCACCACTCA 422
 Db 632 GCAACCCCTCGACCACTTCCAAAGTGTATGGTGAATGTGAGTCTCCGAGTAA 691
 Qy 423 CACTCCCAAGAGAGTGGATGAGCACTGTTTAACTGACTGATCCGCCGTGA 482
 Db 692 TTCCATCCAAACAGCTGATGAATGATGGTTTAACTGACGTCAAGTTCCCAAGA 751
 Qy 483 TTGAGATATCCAGTGGCTTA---TGGCTCTTACAGGTATGACCAATCCCTTCGT 539
 Db 752 TTGAGATATTCAGTGGCTTACGCTTACGCTATGATGACCAAGCAAGAAAACTT 811
 Qy 540 TGAAGTCCAGCAAGAAAGATGTTGCTGGGTTGTGAGTACATGAAACCTGAGCA 599
 Db 812 CACCATCCCAAGAAAGATTAAGTATGTTGCTGATGATGATGATGATGATGATGAT 871
 Qy 600 TGCCAGAGTCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 659
 Db 872 CAAAGCTTACAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 931
 Qy 660 AGCATTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
 Db 932 GCACTTCAATTAACCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 991
 Qy 720 TTAATCTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 779
 Db 992 CTACCTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1051
 Qy 780 TTTTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
 Db 1052 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1111
 Qy 840 TATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 899
 Db 1112 TATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1171
 Qy 900 GATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 959
 Db 1172 TCACCTTAATTCCTGACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1231
 Qy 960 GATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1019
 Db 1232 ACATTTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1291
 Qy 1020 TGTGAAAGGATCAAGAAATTAAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1079
 Db 1292 TCTTAAAGGATTAAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1351
 Qy 1080 A 1080
 Db 1352 A 1352

RESULT 12
 AB023628
 LOCUS AB023628 1832 bp DNA linear VRT 02-APR-1999
 DEFINITION Danio rerio zFT2 gene for alpha(1,3)fucosyltransferase, complete
 cds
 ACCESSION AB023628
 VERSION AB023628.1 GI:4587297
 KEYWORDS alpha(1,3)fucosyltransferase.
 SOURCE Danio rerio DNA.
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE
 AUTHORS Kageyama, N., Natsuka, S. and Hase, S.
 TITLE Molecular cloning and characterization of two zebrafish
 alpha(1,3)fucosyltransferase genes developmentally regulated in

JOURNAL embryogenesis
 REFERENCE J. Biochem. (1999) In press
 2 (bases 1 to 1832)
 AUTHORS Kageyama, N., Natsuka, S. and Hase, S.
 TITLE Direct Submission
 JOURNAL Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate
 School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034,
 Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, Tel: 81-6-6850-5381,
 Fax: 81-6-6850-5383)

FEATURES
 source
 1..1832
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 gene
 332..1510
 /gene="zFT2"
 332..1510
 /gene="zFT2"
 CDS
 /codon_start=1
 /product="alpha(1,3)fucosyltransferase"
 /protein_id="BA076707.1"
 /db_xref="GI:4587298"
 /translation="MDKITVMDKIKLAPSKSNOKLIVTFMLISVCIFFVYVNPNT
 TPEKYPTEGCIHANSSEFCIHYLKNONEICIMKNSYNALEPPAPNPKGODITVL
 IMAPPRVPEPDLDCGLEYNHICQULYRSLSIQKAGHWFHHRDSTDFPPRPF
 OKWIMNMSPTCHNQNDFLNDRENTSSYRSDIPIVYGRVATEEOKFTIPK
 DKIVCWIVSNFOGHHKHFHFEFAKHIVSYGSHNRIINDYDGVNVSCKEYLS
 FENSIRDFEYERLEFNPALGTVPVYIGPPRYNVEEFTIPDAFIHVDDPPTPELADH
 LISLDNEQYKQFPMWRKLIVYSKSPLEHSRCICDYLKRNKHIIAVTDLKGFHWG"

BASE COUNT 585 a 394 c 317 g 536 t
 ORIGIN
 Query Match 21.7%; Score 234.6; DB 5; Length 1832;
 Best local Similarity 55.5%; Pred. No. 4.9e-50;
 Matches 500; Conservative 0; Mismatches 389; Indels 12; Gaps 2;

Qy 183 TAATGAATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
 Db 619 TCAAGACATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
 Qy 243 CTGCCAAGCAATGTTCAACATCCCAAGATGCTCATTCACACAGCAGCAGCAGCAGCAGCAGC 302
 Db 679 TTGGTGTGGAGTATTAATCAATCCAGGCTGTCAATTAATCAATCAATCAATCAATCAATCAAT 738
 Qy 303 CAATCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
 Db 739 GAAAGCCATGAGATTAATGTTTACCAACAGATATGAGC-----ACAGATTTTCC 789
 Qy 363 TCAGCAAGTACGACCCCTTCAGAAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 422
 Db 790 GCAACCCCGCGACAGATTTCCAAAGTGTGATATGATGATGATGATGATGATGATGATGATGATGAT 849
 Qy 423 CACTCCCAAG 482
 Db 850 TTGCAATCAAAATGACTTCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
 Qy 483 TTGAGATATCCAGTGGCTTATGCTTCTGACGATGACCA---AATCCCTTCGTT 539
 Db 910 TTGAGATATTCAGTGGCTTATGCTTCTGACGATGACGATGACGATGACGATGACGATGACGATGAC 969
 Qy 540 TGAAGTCCAGCAAGAAAGATTTGCTGGGTTGTGAGTACTGAAACCTGAGCA 599
 Db 970 CACCATCCCAAGAAAGATTAAGTATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1029
 Qy 600 TGCCAGAGTCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
 Db 1030 CAAAGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
 Qy 660 AGCATTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
 Db 1090 GCACTTCAATTAACCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1149
 Qy 720 TTAATCTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 779

Db 1150 CTACCTGTCCTCTAGAACTCCATTCACAGAGACTATTCTACAGGAAAGCTTCCATCC 1209

QY 780 TTTTCTGGCTGGCTCTGTACCTGTGTCTCTGGACCATCTAGGGAAACATATAGAAATTA 839

Db 1210 TTTAGACACTCGTACAAATTCCTGTGTATTGGCCCCCCCAAGATATTACTGTAGAAAGTTT 1269

QY 840 TATTCAGCAATTCATCTATCTATGTAGAGATTATTAACCTCCAGCTAGCGAATA 899

Db 1270 CATACCAAGAGCCGCTCATCCATGTGGATGATTGCCAACCTCCCAAGAACCTGGGGGA 1329

QY 900 GTATCTGAAGGAATCGACAAACAAATTAATTTTACCTTAGTACTTAACTGTGAGGAA 959

Db 1330 TCACCTTATATCTTGGACCAAAAGAGGAACAGTACAAACAGTTTTTCAACTGAGAGAA 1389

QY 960 GGATTTCACTGTAATCTTCCACGATTTGGGAAATCCATCACTGTTGGCTTGCAGTCA 1019

Db 1390 ACTTTTGTGTTCCAAAGAGTACATCGCTTTGGGCTTGAACATCTCCTTAGAGATTTGTGATTA 1449

QY 1020 TGTGAAAGGATCAAAATATATAGCTGTGTGTAATTTGAGAAATGTTTGGAAATTA 1079

Db 1450 TCTTAAAGGAATTAAGCACTATATAGCCGTTTAGCGATCTGAAAGAGGTGTTCTGTGGGTTA 1509

QY 1080 A 1080

Db 1510 A 1510

RESULT 13			
AB035907			
LOCUS	AB035907	1092 bp	DNA linear
DEFINITION	Xenopus laevis gene for xft1, complete cds.		
Accession	AB035907		
			VRI 05-JAN-2002

SOURCE Xenopus laevis bone marrow DNA.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.

REFERENCE	AUTHORS	TITLE	JOURNAL
1	(sites)		
	Kaneko, M., Nishihara, S., Kitano, T., Narimatsu, H. and Saitou, N.	The evolutionary history of glycosyltransferase genes	Unpublished

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 1092)
Kaneko, M., Saitou, N. and Kitano, T.
Direct Submission
Submitted (17-DEC-1999) Miki Kaneko, National Institute of

JOURNAL
Submitted (17-DEC-1999) Mika Kameo, National Institute of
Genetics, Laboratory of Evolutionary Genetics, Yata 1111, Mishima
Shizuoka 411-8540, Japan (E-mail: mikakameo@med.id.yamagata-u.ac.jp,
Tel: 81-559-81-6790, Fax: 81-559-81-6789)
To: journal@nsg.riken.go.jp

EMPLOYED	LOCATION/QUALITIES
SOURCE	1. .1092

```

"organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_type="bone marrow"
gene
1. .1092
/gene="xft1"
CDS
1. .1092

```

CDS

BASE COUNT
ORIGIN

256 a	281 c	294 g	261 t
-------	-------	-------	-------

Query Match	20.6%	Score 222.6;	DB 5;	Length 1092;
Best Local Similarity	57.3%;	Pred. No. 6.3e-47;		
Matches 484; Conservative	0;	Mismatches 349;	Indels 12;	Gaps 4.

OY	184	AATGAAACTATATCTGGTGTGGGTGGGACATTTGGGACACACCTTTGACCTTACATCC	243
Db	199	AAGCAAGTCACTGTCTTAAATATGATGTAGGCATTTGGCAAAAGCGGCTTGGGAT	258
OY	244	TCCCAACACATGTGCAACATCCAAACGAGATCCCATCTCCAAACGAGACCTTACGTACAC	303
Db	259	TCCACATGCTCTTCAATGATCAGGGGTCCACATCCAAACCAAGAGATCTTTACAG	318
OY	304	AAATCCATGCAAGTTCTGTATCATCACGAGACATCAGTTGGGATGTGACAAATTTACT	363
Db	319	GAGGAGAACCCCAATATATGTTCCACACCGGACAT--TTCTGATTTCAATGACCTTCTT	375
OY	364	CAGCAAGCTAGGCCACCTTCCAGAAATTTGATGATATTTGGAAATGACCAACTAC	423
Db	376	TTCCGAAGGGCCAGAGCTCCCAAAATGTGATCTGATGACTTGAAATCTCCGTCAC	435
OY	424	ACTGCCAAAAGGTGATGTAGACACTGTGTTTACCTGACGTGACGTACGTCCGCGGAT	483
Db	436	TGGCTCTGCTGGCAGATCTGGGGGACATCTTCACTGGACCAATGTCTCTACAGGTGAC	495
OY	484	TCAGATATCCAGTGCCTTATGGCTTGTGAGGTTAAGCAACAATCCCTTCTGTGTTGAA	543
Db	496	TGCGATATCTTCAATGCGCTACGGCTTCTT--ATTCTCCAAACAAAGTCCAAATATCTC	552
OY	544	GTGCCAAGCAAGAAATTTGGTGTCTGGGTGTGAGTACGTGAAACCTGAGATGCC	603
Db	553	CTGCCGGGCAAGAAAGCTGTGGGCTGGGTGATCAGCACTGGAAAGAGAGCCACGA	612
OY	604	AACGTCAAGTATACATGACATCAGCAAAACGATGAAATGCACTACCTACGCGGCAACA	663
Db	613	AGGGTCAAGTATTCAGACGACTGAGAACTACGTGAATGTGAGCTTTACGGCC--GG	669
OY	664	TTTGGAGATATGTCAATGATTAATAAATTTGATCTTACCAATATCTGTGTAAATTTAT	723
Db	670	TATGGCGTGATTTAAAAAGAGCAACATCGTCAAGACGGCTCTCGGATTTAAATTTTAC	729
OY	724	CTTTCCTTTAAATATCAATCACAAGATTTCAATCAGGAAACATAT--ACAATCT	780
Db	730	TTGGCTTTTGAATTTCTTCCACACAGGATTTACATCAGGAAACGTGTGGAAGAACCG	789
OY	781	TTTTCGGCTGGCTGTACGTGTGTTCTGGGACACATCTAAGGAAACTGTGAGAAATAT	840
Db	790	TTTAAATTCAAACCATCCCCATGTGATGAGGGCCACAGCCGCTATTAATCTACGAGATGTC	849
OY	841	ATTCCAGAGATTCATTCATTCATGTGGAGATTAATACTCTCCAGTGAAGTACGAAG	900
Db	850	ATACCGCGAGCTCTTTCATTCAGGTGGACAGATTTCTCCAGGCCACGAGAAAGCTGGCCATG	909
OY	901	TATCTGAGAGAAATGCAAAAACAAATAGTTATTAACCTTAGTTAACTGAGAGAG	960
Db	910	TACCTGAAGCACTGTGATTAATAAATTTACTCTGTACCGAGGATTTTACGTGAAGAG	969
OY	961	GATTTTCACTGAAATCTTCCAGATTTTGGGATCAATGACAAATGTTGGCTGCGCATAT	1020
Db	970	CGTTACAGCTGACGTACGTCTTCTTGGAGCAGAGATTTACTGCAACCGCGTCCAGAGC	1029
OY	1021	GTGAA	1025
Db	1030	GTCAA	1034

RESULT 14				
AF288369	AF288369	1264 bp	DNA	linear
LOCUS				
DEFINITION	Gallus gallus alpha (1,3) lucosyltransferase gene,			VRT_02-Dec-2001
ACCESSION	AF288369			complete cds.
VERSION	AF288369.1	GI:17224489		
KEYWORDS				

SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1264)
AUTHORS Foley, A.C., Schmitz, B., Stern, C.D. and Strell, A.
TITLE Expression of fucosyltransferases in chick embryos
JOURNAL Unpublished
2 (bases 1 to 1264)
REFERENCE Foley, A.C., Schmitz, B., Stern, C.D. and Strell, A.
AUTHORS Direct Submission
JOURNAL Submitted (19-JUL-2000) Genetics and Development, Columbia
University, 701 West 168th Street, New York, NY 10032, USA
FEATURES
source
1..1264
Location/Qualifiers
/organism="Gallus gallus"
/db_xref="taxon:9031"
CDS
1..1062
/note="fuc"
/codon_start=1
/product="alpha (1,3) fucosyltransferase"
/protein_id="AA136994.1"
/db_xref="GI:17224490"
/translation="SMKLLTFLSFISFSLFSLRQFRPEPDYNSYSTERLP
AGSTMPLKPSERMLTILMTWPFGRHFNFSNCESELDPNCHPTNRSWSQADAVIM
HHRDYMDRELRSLRPLRPSORIMFNMSPSHSMDLFLNLMYRSDIFL
PYGELQLSEPPQFSPKTKMVAWVSWMRSPSHKYEELKRIIPDYGRHLP
LSMDLHDSOYNFYLAENSEHDKYIEKMRNLSSGTLIPVLYGRIYEHPLP
PDSFIHVDPSAGELAOYLQLSRDTKRYRFPQWRRLKYLVLGIMAMQYCRACHF
LQRTASIVVYNLSKWEY"
BASE COUNT 286 a 355 c 319 g 304 t
ORIGIN
Query Match 20.4%; Score 220.2; DB 5; Length 1264;
Best Local Similarity 58.0%; Pred. No. 2.6e-46;
Matches 447; Conservative 0; Mismatches 318; Indels 6; Gaps 3;

OR 733 GAAATTCATCCACAGAGATTACATCCGAGAAAGCTATAC--AAATCTTTCTGGGCT 789
DB 712 GAGAACCTCAGAGACCAAAAGACTACATCAGAGAGCTCGAGATGCCCTGCTCT 771
OR 790 GGCTCTGCTACCTTTGTTCTGGGACCATCTAGGAAACATATAGATTTATTCACGA 849
DB 772 GGACACATCCCACTGTTCTGGGGCCACTCGAGAAACATATAGACTTTCTGGCCCC 831
OR 850 GATTCATTCATTCATGGAAGATTATACCTCCAGTAGTAGCAAGTATGAG 909
DB 832 GACTCTTCATCCACGTCGATGACTTGGCAGTCTGCTGTAATAGCGGCACTACCTG 891
OR 910 GAAATTCATCCACAGAGATTATACCTTTAGTACTTTACTGAGAGAG 960
DB 892 GAATGAGCAGAGGACACCGAGAGTACACCGCTACTTCAGTGGCGCAGG 942
RESULT 15
GGU73678 2657 bp DNA linear VRT 20-DEC-1996
LOCUS Gallus gallus alpha-1,3-fucosyltransferase (CFT1) gene, complete
DEFINITION cds.
ACCESSION U73678
VERSION U73678.1 GI:1657998
KEYWORDS chicken.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2657)
AUTHORS Lee, K.P., Carlson, L.M., Woodcock, J.B., Ramachandra, N.,
Schultz, T.L., Davis, T.A., Lowe, J.B., Thompson, C.B. and Larsen, R.D.
TITLE Molecular cloning and characterization of CFT1, a developmentally
regulated avian alpha(1,3)-fucosyltransferase gene
JOURNAL J. Biol. Chem. 271 (51), 32960-32967 (1996)
MEDLINE 97115837
REFERENCE 2 (bases 1 to 2657)
AUTHORS Lee, K.P., Carlson, L.M. and Larsen, R.D.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1996) Immune Cell Biology Program, Naval Medical
Research Institute, 8901 Wisconsin Ave., Bethesda, MD 20889, USA
FEATURES
source
1..2657
Location/Qualifiers
/organism="Gallus gallus"
/db_xref="taxon:9031"
gene
170..1240
/gene="CFT1"
CDS
170..1240
/gene="CFT1"
/note="similar to mammalian fucosyltransferase FucTIV"
/product="alpha-1,3-fucosyltransferase"
/protein_id="AAC60060.1"
/db_xref="GI:1657998"
/translation="MELGRMSPAPRPGCPRRMRMRWALLGALALAYCYVRELR
RGSAGREGEVTVLWMEPGRPWRPADCRVRVITGCLISADGRGEARVLEFH
HDDIALHGRQGLPRGPPRPDRPDRWMMWSPSPSGRLGAGLEFNMYSTRSDY
FVPLGTYLPPSPRPVPLPRKSLVAMVTSNNNEHARVRYRQLKEHLIPDYGARG
MLBGSVAVTKSAKLYLAFENSOHTDVTETKLMWNAASAVPVGLPRAANERF
IPADSFTHVDPSRLATVLEKFLDKNPSYRFRYFAMWKNKEVHTSPWDHYCKVC
EAVRTAGNOLKTQVNLGWFES"
BASE COUNT 581 a 712 c 818 g 546 t
ORIGIN
Query Match 16.3%; Score 176.2; DB 5; Length 2657;
Best Local Similarity 53.1%; Pred. No. 6.1e-35;
Matches 451; Conservative 0; Mismatches 383; Indels 15; Gaps 3;

Tue Oct 8 10:18:26 2002

us-09-744-748-4_copy_1289_2368.rge

Page 16

```
OY 253 ATGTTCAATCATCAAGATGCCATCTCACAAGGACGGTTCACTGATACAAATCCCAT 312
Db 401 CGGTACAAATCATCGGGCTGCTCTCTAGAGCGCGGGGCTACGGCGAGGCTCGG 460
OY 313 GCAGTCTCATGATTCACGAGACATCACTGGGATCTGACAATTTTACCTCAGCAAG-- 370
Db 461 GCGGTCTCTTCACACACCGCGACCTGGCGCTGCAAGCGCGCCAGGGGCTGCCCGCGG 520
OY 371 -----CTAGGCCACCTCTCCAGAAATGATTTGATGAATTTGAAATCACCACATC 423
Db 521 CCCCCCGCCGACCCCGCGGCGAGCGGGTGGTGAATGAACCTGAGTCCGCTCGGCAC 580
OY 424 ACTCCCAAAAGAGTGGCATTTGAGCACTGTTTAACTGACTCTGACTTACCGCGGTGAT 483
Db 581 TCCCGGGGCTCGGGGCTGGCGGCTCTTCAACTGAGACCATGCTGATCCGAGCGGAC 640
OY 484 TCAGATATCCAAAGTCCCTTATGGCTTTTGAACGCTAGCACAAATCCCTGCTTGA 543
Db 641 TCGGAGCTCTTCGCTCCCTACGGTACTCTACGAGCCGCGCTGCGGCCCTTGCTG 700
OY 544 GTGCCAAGCAAGAGAAATTTGGTGTGCTGGGTTGTGATTAAGTGAACCTGAGCATGCC 603
Db 701 CTGCTTCGCAAAAGCGCGCTGTGGCTGTGGTCACTACGACCACTGGAACGAGGAGCGCC 760
OY 604 AGAGTCAAGTATTTACATAGCTAAGCAAAAGCATTTGAATTCATACCTAGCGGCAAGCA 663
Db 761 GCGGTGCGCTACTACCGGCGAGCTGAAGGAGCACCTGCCATCGACGTGTACGGG---GCG 817
OY 664 TTTGAGATATGTCAATGATTAATAAATTTGATTCCTACATATCGTGTGTAATTTAT 723
Db 818 CGCGGATGCGGCTGCTGTGAGGAGCGAGCTGTGAAGAGCGTGTGCGCTACAAAGTTCTAC 877
OY 724 CTTTCTTTGAAAAATTCATCCACAAGATTTACATCAAGGAAAGCTAT---ACAATGCT 780
Db 878 CTGGCTTTCGAGAACTCCACAGCAGGACTATACACCGAGAGCTGTGGAAGAACGCC 937
OY 781 TTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 938 TTTGCGCGGCGAGCGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
OY 841 ATTCAGAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
Db 998 ATCCCGCGGCGACTCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1057
OY 901 TATCTGAAGAGTGCACAAAACAATAAGTTATACCTAGTACTTAAGTGAAGAG 960
Db 1058 TACCTGAAGTTCTCGATTAACAAAACAAGCCAGCTACAGAGGTATTTGCTGCGGAGC 1117
OY 961 GATTTCAGTGAATCTTCACAGATTTGGCAATTCACATGATGTTGGCTTGGCATCAT 1020
Db 1118 AAGTATGATGCCAGTCACGTCCTTTCTGAGATGAGCATTAAGGTTTGGAGGCC 1177
OY 1021 GTGAAAGG 1029
Db 1178 GTGAGGAGG 1186
```

Search completed: October 6, 2002, 04:35:48
Job time : 1497.36 secs

PI Narimatsu H, Kudo T, Sasaki K;
 XX WPI: 2000-183120/16.
 DR P-PSDB: AAB0996.
 XX
 PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT epitope and an antibody recognizing it useful for diagnosis of brain
 PT and kidney disease and cancer.
 XX
 PS Claim 4: Page 143-150; 172pp; Japanese.
 XX
 CC The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used
 CC for the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequences
 CC AA292646 and AA292647 represent cDNAs encoding human
 CC alpha-1,3-fucosyltransferase (AA80996).
 CC
 XX
 SO Sequence 2676 BP; 819 A; 527 C; 476 G; 854 T; 0 other;
 Query Match 100.0%; Score 1080; DB 21; Length 2676;
 Best Local Similarity 100.0%; Pred. No. 4e-300;
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 715 CACACCCCCCAAGAGGCGATGAGCACTGTTAACTGACTGACTTCCGCCCT 774
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 481 GATTCGATATCCCAAGTCCCTATGCGCTTCTGACGTAAGCAACAATCCCTCGTGT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 775 GATTGAGATATCCAGTCCCTATGCGCTTCTGACGTAAGCAACAATCCCTCGTGT 834
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 541 GAAGTGCACCAAGCAAGAAATGTTGCTGCGGTGTGATGATGAGAACCCCTGAGAT 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 835 GAAGTGCACCAAGCAAGAAATGTTGCTGCGGTGTGATGATGAGAACCCCTGAGAT 894
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 601 GCCAAGTCAAGTATATTAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 895 GCCAAGTCAAGTATATTAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 954
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 661 GCATTTGGAGAAATATGCAATGATTAATAATTTGATTCCTACCAATATCTGTAATTT 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 955 GCATTTGGAGAAATATGCAATGATTAATAATTTGATTCCTACCAATATCTGTAATTT 1014
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 721 TATCTTCTCTTGAATAATTCATCCAGAGATTCATCCAGCAAGCAAGCAAGCAAG 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1015 TATCTTCTCTTGAATAATTCATCCAGAGATTCATCCAGCAAGCAAGCAAGCAAG 1074
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 781 TTTCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1075 TTTCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 841 ATTCAGAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1135 ATTCAGAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1194
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 901 TATTCGAAGAGAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1195 TATTCGAAGAGAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1254
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 961 GATTCACCTGTAATCTTCACAGATTTGGGAATCAGATGATTTGGCTTGCATCAT 1020
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1255 GATTCACCTGTAATCTTCACAGATTTGGGAATCAGATGATTTGGCTTGCATCAT 1314
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1021 GTGAAAAGGCATCAAGAAATATAGCTGTTGTAATTTAGAGAAATGTTGGAAATTA 1080
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1315 GTGAAAAGGCATCAAGAAATATAGCTGTTGTAATTTAGAGAAATGTTGGAAATTA 1374
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2
 AA292646
 ID AA292646 standard; cDNA; 2822 BP.
 XX
 AC AA292646;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:4.
 XX
 KW Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
 KW treatment; cancer; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS /tag= a
 FT /product= "Human alpha-1,3-fucosyltransferase"
 FT
 FT
 FT
 FT
 PD 10-FEB-2000.
 XX
 PD 29-JUL-1999; 99WO-JP04092.
 XX
 PF 29-JUL-1998; 98JP-0213823.
 XX
 PR

XX (KYOM) KYOMA HAKKO KOGYO KK.
XX Naimatsu H, Kudo T, Sasaki K;
XX WPI: 2000-183120/16.
XX P-PSDB: AAY80996.
XX Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
XX epitope and an antibody recognizing it useful for diagnosis of brain
XX and kidney disease and cancer.
XX
XX Claim 4: Page 134-142; 172pp; Japanese.
XX
XX The invention relates to a novel alpha-1,3-fucosyltransferase which
XX transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
XX (generating the Lewis x or y epitope). It does not transfer a fucose
XX moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
XX and therefore does not generate the sialyl-Lewis x epitope. The
XX invention also relates to DNA sequences encoding alpha-1,3-
XX fucosyltransferase and expression vectors and host cells comprising
XX these DNA sequences. The invention additionally encompasses the
XX preparation of alpha-1,3-fucosyltransferase via the culture of
XX transformed cells or by expression of the protein in a transgenic animal;
XX antibodies which recognise alpha-1,3-fucosyltransferase; methods for
XX screening potential inhibitors or potentiators of
XX alpha-1,3-fucosyltransferase activity or expression; the preparation of
XX compounds having fucose-containing sugar chains by use of the protein;
XX and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
XX Alpha-1,3-fucosyltransferase has a similar substrate range to the known
XX FUC-TIV and is expressed mainly in brain and kidney tissues.
XX Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
XX potentiators and inhibitors may be used for the treatment and diagnosis
XX of diseases of the brain and kidney, and of cancers. They may be used for
XX the identification of substances which affect the activity or expression
XX of alpha-1,3-fucosyltransferase: such substances may be used
XX therapeutically. The knockout animals can be used to study the mechanisms
XX of action and expression of alpha-1,3-fucosyltransferase. Sequences
XX alpha-1,3-fucosyltransferase (AAY80996).
XX
SQ Sequence 2822 BP: 918 A; 485 C; 477 G; 942 T; 0 other:
Query Match 100.0%; Score 1080; DB 21; Length 2822;
Best Local Similarity 100.0%; Pred. No. 4,1e-300;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACATCAACATCCAAAGAAATCTTGGCCATTTTAAATGTCTGCATTAATCCCTGGGC 60
DB 1289 ATGACATCAACATCCAAAGAAATCTTGGCCATTTTAAATGTCTGCATTAATCCCTGGGC 1348
QY 61 TGTTCATGCGATGCTCTCATTTACATCAACACCTACCAAGAGTGTGATCTTCACTGCA 120
DB 1349 TGTTCATGCGATGCTCTCATTTACATCAACACCTACCAAGAGTGTGATCTTCACTGCA 1408
QY 121 ATGGAATCAAGCAGCTCTGTGCTGAATAATGAAAACTTTTCCACCAAAACTGATAT 180
DB 1409 ATGGAATCAAGCAGCTCTGTGCTGAATAATGAAAACTTTTCCACCAAAACTGATAT 1468
QY 181 TTTAATGAATACATATCTGTGTGTGGTGGCCATTTGGGAGACCTTTGACCTTA 240
DB 1469 TTTAATGAATACATATCTGTGTGTGGTGGCCATTTGGGAGACCTTTGACCTTA 1528
QY 241 TCCGCGCAAGCATGTTTCAACATCCAAAGATGCCATCTACCAAGGACGTTCACTGTAC 300
DB 1529 TCCGCGCAAGCATGTTTCAACATCCAAAGATGCCATCTACCAAGGACGTTCACTGTAC 1588
QY 301 AACAAATCCCATGCACTGTGATCATCAACCGAGACATGATGGATCTGACAAATTTA 360
DB 1589 AACAAATCCCATGCACTGTGATCATCAACCGAGACATGATGGATCTGACAAATTTA 1648
QY 361 CCTCAGCAAGCTAGGCGACCTTCCGAAATGATTTGGATGATTTGGAATACCAACT 420
|||||

DB 1649 CCTCAGCAAGCTAGGCGACCTTCCGAAATGATTTGGATGATTTGGAATACCAACT 1708
QY 421 CACACTCCCAAAAGAGTGGCATTTGAGACCTGTTTAACTGACTGTGACTACCGCGT 480
DB 1709 CACACTCCCAAAAGAGTGGCATTTGAGACCTGTTTAACTGACTGTGACTACCGCGT 1768
QY 481 GATTCAGATATCCAGTGGCTTATGCTTCTTGACGGTAAAGCAAAATCCCTTGTGTT 540
DB 1769 GATTCAGATATCCAGTGGCTTATGCTTCTTGACGGTAAAGCAAAATCCCTTGTGTT 1828
QY 541 GAAATGCCAAGCAAAAGAAATGCTGTGCTGGTGTGAGTAACGTAACCTGTGAGAT 600
DB 1829 GAAATGCCAAGCAAAAGAAATGCTGTGCTGGTGTGAGTAACGTAACCTGTGAGAT 1888
QY 601 GCCAGATCAAGTATTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 1889 GCCAGATCAAGTATTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 1948
QY 661 GCATTTGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 1949 GCATTTGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2008
QY 721 TATCTTCCCTTGAATAATTCATCCCAAGGATTCATCAGGAAAGCATATCAATGCT 780
DB 2009 TATCTTCCCTTGAATAATTCATCCCAAGGATTCATCAGGAAAGCATATCAATGCT 2068
QY 781 TTTCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 2069 TTTCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2128
QY 841 ATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 900
DB 2129 ATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2188
QY 901 TATCTGAGAGAGTGCACAAAACATTAAGTATTAATCTTACCTTAACTGAGAGAG 960
DB 2189 TATCTGAGAGAGTGCACAAAACATTAAGTATTAATCTTACCTTAACTGAGAGAG 2248
QY 961 GATTCACATGTAATCTTCCAGATTTTGGGAATCAGATGATGATGATGATGATGAT 1020
DB 2249 GATTCACATGTAATCTTCCAGATTTTGGGAATCAGATGATGATGATGATGATGAT 2308
QY 1021 GTGAAAAGCATCAGAAATTAAGTGTGTTGTAATTAAGAAATGTTTGAATTA 1080
DB 2309 GTGAAAAGCATCAGAAATTAAGTGTGTTGTAATTAAGAAATGTTTGAATTA 2368
RESULT 3
AA292672
ID AA292672 standard; cDNA; 2036 BP.
XX
XX AA292672;
AC
XX
XX 05-JUN-2000 (first entry)
DE Murine alpha-1,3-fucosyltransferase gene exon 1 DNA, SEQ ID NO:30.
XX
XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KW brain; kidney; recombinant expression; transgenic animal; knockout
KW animal; FUC-TIV; drug screening; inhibitor; potentiators; diagnosis;
KW treatment; cancer; murine; mouse; ss.
XX
XX Mus sp.
OS
XX
XX WO200006708-A1.
PN
XX
XX 10-FEB-2000.
PD
XX
XX 29-JUL-1999; 99WO-JP04092.
PF
XX
XX 29-JUL-1998; 98JP-0213823.
PR
XX
XX (KYOM) KYOMA HAKKO KOGYO KK.
PA

XX Narimatsu H, Kudo T, Sasaki K;
PI
XX WPI; 2000-183120/16.
DR

PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer. -

Example 8; Page 162-164; 172pp; Japanese.

The invention relates to a novel alpha-1,3-fucosyltransferase which transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglucosamine (generating the Lewis x or y epitope). It does not transfer a fucose moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglucosamine and therefore does not generate the sialyl-Lewis x epitope. The invention also relates to DNA sequences encoding alpha-1,3-fucosyltransferase and expression vectors and host cells comprising these DNA sequences. The invention additionally encompasses the preparation of alpha-1,3-fucosyltransferase via the culture of transformed cells or by expression of the protein in a transgenic animal; antibodies which recognise alpha-1,3-fucosyltransferase; methods for screening potential inhibitors or potentiators of alpha-1,3-fucosyltransferase activity or expression; the preparation of compounds having fucose-containing sugar chains by use of the protein; and knockout non-human animals lacking alpha-1,3-fucosyltransferase. Alpha-1,3-fucosyltransferase has a similar substrate range to the known FUC-TV and is expressed mainly in brain and kidney tissues. Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies, potentiators and inhibitors may be used for the treatment and diagnosis of diseases of the brain and kidney, and of cancers. They may be used for the identification of substances which affect the activity or expression of alpha-1,3-fucosyltransferase; such substances may be used therapeutically. The knockout animal can be used to study the mechanisms of action and expression of alpha-1,3-fucosyltransferase. Sequence AA952845 represents cDNA encoding murine alpha-1,3-fucosyltransferase (AA980995), and sequences AA925670-292673 are murine alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid amplification of cDNA ends).

SQ Sequence 2036 BP; 666 A; 365 C; 348 G; 657 T; 0 other;

Query Match	86.2%;	Score 931.2;	DB 21;	Length 2036;
Best Local Similarity	91.4%;	Pred. No. 2.2e-257;		
Matches 987; Conservative	0;	Mismatches 93;	Indels 0;	Gaps 0

Qy	1	ATGCAATCAACATCCAAAGAAATTCCTGGCCATTTTTAATGTGTGCATTTACCTGGGC	60
Db	9	ATGCAATCAACATCCAAAGCAATTCCTGGCCATTTTTAATGTGTGCATTTACCTGGGC	68
Qy	61	TGTTTATGSCAATGCTTCTCTCAATTTAACTCAAAACATCCAAAGCGTGAATCTGATGCCA	120
Db	69	TGCTCTATGGCAATGCTGTCTCTATTATATATCAAGCCCAACAGCTGGGCTCTCAATCCA	128
Qy	121	ATGGAAATCAAGCAGCTGCTGCTGAAAAATGAAAAATCTCTTTTCCAAACCAAACTGATTAT	180
Db	129	ATGAGATCTGCAAGTTCTGTGTCTGAAATTAAGAAAAATTTCTTCTCCAAATAAACTGATTAT	188
Qy	181	TTTAAATGAACATCATTTCTGTGTGTGGGTGTGGCCATTTTGGCAGACCTTTTACCTTACA	240
Db	189	TTTAAACCAATCAATCAATTCCTGTGTGTGGGTATATGGCCATTTTGGGAGACCTTTTACCTTACA	248
Qy	241	TCTCTGCACGAATGTTCAACATCCAAAGATGTCATCTCAACAGGACGGTTCACCTGTAC	300
Db	249	TCTCTGCACGAATGTTCAATATATCCAAAGGGGTCATCTCAACAAACAGCGCTCATTTGTAC	308
Qy	301	AACAAATTCGCATGCAAGTTCTGTGATCATCCAGACATCAGTGTGGGATCTGACAAATTTTA	360
Db	309	AACAAATTCGCATGCGGCTCTATTCACCATAGAGATCATCAGCTGGGATCTGACATACTTA	368
Qy	361	CCTCAGCAAGTATGGCGCAACCCCTTCAGAAATGTGATTTGGATTTGGAAATCAACCACT	420
Db	369	CCTCAGCAAGTATGGCGCAACCCCTTCAGAAATGTGATTTGGATTTGGAAATCAACCACT	428

OY	421	CACACGCCCAAAAGAGTGGCATTTGAGCACTTGTTAACTACGCTGACTCTTCACCCGCT	480
Db	429	CACACCCCAAAAGAGTGGCATTTGAGCACTTGTTAACTACGCTGACTCTTCACCCGCT	488
OY	481	GATTCAGATATCCAGTAGTGGCTTATATGGCTCTCTTGAGCGGTAGCACAAATCCCTTCGTGTT	540
Db	489	GATTCAGATATCAAGTGGCTTATATGGCTCTCTTGAGCGGTAGCACAAATCCCTTCGTGTT	548
OY	541	GAAGTCCCAAGCAAGAGAAATTTGGTGTGCTGGTGTGAGTTAACTGSAACCCCTGAGCAT	600
Db	549	GAAGTCCCAAGCAAGAGAAATTTGGTGTGCTGGTGTGAGTTAACTGSAACCCCTGAGCAT	608
OY	601	GCCAGAGTCACATATTTACAAATGAGCTAACCAAAAGCATTTGAAATCCATACCTAGCGCAA	660
Db	609	GCCAGAGTCACATATTTACAAAGAGCTAGCAAGATATTTGAAATCCACACTATAGCGCAA	668
OY	661	GCAATTTGAGAAATATGTCACATGATPAAAAATTTGATTTCCACCATATCGCTGTGTAATTT	720
Db	669	GCAATTTGAGAAATATGTCACATGATPAAAAATTTGATTTCCACCATATCTACTTTGTAATTT	728
OY	721	TATCTTTCTTTGAAATATTCATTCACAGGATTTACATCAGGAAAAAGCTATATCAATGCT	780
Db	729	TATCTTTCTTTGAAATATTCATATTCACAAAGATTTACATCAGGAAAAAGCTATATCAATGCT	788
OY	781	TTTTTGCGCTGCTGTACCTGTGTGTTCTGGACCATCTAGGAAAACTATGAAATTTAT	840
Db	789	TTTTTGCGCTGCTGTACCTGTGTGCTCGGTCATCTAGGAAAACTATGAAATTTAT	848
OY	841	ATTCCAGCAGATTCATTCATTCATGATGGAAGATTTAACTCTCCACAGTAGCATGCAAG	900
Db	849	ATTCCAGCAGATTCATTCATTCATGATGGAAGATTTAACTCTCCACAGTAGCATGCAAAA	908
OY	901	TATCTCAAGAGATGAGCAAAAAACATTAAGTTATTAAGCTAGTTACTTTAACTGAGAGAG	960
Db	909	TATCTCAAGAGATGAGCAAAAAACATTAAGTTATTAAGCTAGTTAACTGAGAGAG	968
OY	961	GATTTACGCTTAAACTTCCACAGATTTGGGAAATCAGATCATGATGTTGGCTTGCGACTAT	1020
Db	969	GATTTACGCTTAAACTTCCACAGGTTTTTGGGAAATCAGATCATGATGCTGGGACTGAGCAAT	1028
OY	1021	GTCAAAAGCATCAAGATATTAAGTCTGTTGGTATATTTAGAGAAATGTTTTGGAATTTAA	1080
Db	1029	GTCAAAAGCATCAAGATATTAAGTCTGTTGGTATATTTAGAGAAATGTTTTGGAATTTAA	1088

RESULT	4
AAZ92645	
ID	AAZ92645 standard; cDNA; 2170 BP.

DT 05-JUN-2000 (first entry)

DE Murine alpha-1,3-fucosyltransferase cDNA.

KM Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope
KM brain; recombinant expression; transgenic animal; knockout
KM animal; Fuc-TIV; drug screening; inhibitor; potentiator; diagnosis
KM treatment; cancer; murine; mouse; ss.

OS Mus sp.

EH	Key	Location/Qualifiers
FT	CDS	115..1194
FT		/tag= a
FT		/product= "Murine alpha-1,3-fucosyltransferase"

PN	WO200006708-A1.
XX	
PD	10-FEB-2000.
XX	
PF	29-JUL-1999; 99WO-JP04092.

XX	29-JUL-1998;	98JP-0213823.	
PR	(KYOM) KYOWA HAKKO KOGYO KK.		
PA	Narimatsu H, Kudo T, Sasaki K;		
PI	WPI: 2000-183120/16.		
XX	P-PSDB; AAH80995.		
DR			
XX			
PT	Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x		
PT	epitope and an antibody recognizing it useful for diagnosis of brain		
PT	and kidney disease and cancer.		
PS	Claim 4; Page 127-134; 172pp; Japanese.		
XX			
CC	The invention relates to a novel alpha-1,3-fucosyltransferase which		
CC	transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine		
CC	(generating the Lewis x or y epitope). It does not transfer a fucose		
CC	moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine		
CC	and therefore does not generate the sialyl-Lewis x epitope. The		
CC	invention also relates to DNA sequences encoding alpha-1,3-		
CC	fucosyltransferase and expression vectors and host cells comprising		
CC	these DNA sequences. The invention additionally encompasses the		
CC	preparation of alpha-1,3-fucosyltransferase via the culture of		
CC	transformed cells or by expression of the protein in a transgenic animal;		
CC	antibodies which recognize alpha-1,3-fucosyltransferase; methods for		
CC	screening potential inhibitors or potentiators of		
CC	alpha-1,3-fucosyltransferase activity or expression; the preparation of		
CC	compounds having fucose-containing sugar chains by use of the protein;		
CC	and knockout non-human animals lacking alpha-1,3-fucosyltransferase.		
CC	Alpha-1,3-fucosyltransferase has a similar substrate range to the known		
CC	FUC-TV and is expressed mainly in brain and kidney tissues.		
CC	Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,		
CC	potentiators and inhibitors may be used for the treatment and diagnosis		
CC	of diseases of the brain and kidney, and of cancers. They may be used for		
CC	the identification of substances which affect the activity or expression		
CC	of alpha-1,3-fucosyltransferase; such substances may be used		
CC	therapeutically. The knockout animals can be used to study the mechanisms		
CC	of action and expression of alpha-1,3-fucosyltransferase. Sequence		
CC	AA92645 represents cDNA encoding murine alpha-1,3-fucosyltransferase		
CC	(AAH80995), and sequences AA92670-92673 are murine		
CC	alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid		
CC	amplification of cDNA ends).		
XX			
XX	Sequence 2170 BP: 715 A: 395 C: 368 G: 692 T: 0 other:		
SO			
	Query Match	86.2%; Score 931.2; DB 21; Length 2170;	
	Best Local Similarity	91.4%; Pred. No. 2.2e-257;	
	Matches 987; Conservative	0; Mismatches 93; Indels 0; Gaps 0	
OY	1 ATGACATCAACATCCAAAGGAAATCTCTGCCCATTTTAAATTGTCTGCATTATCCGGGC	60	
DB	115 ATGACATCAACATCCAAAGGAAATCTCTGCCCATTTTAAATTGTCTGCATTATCCGGGC	174	
OY	61 TGTTCATGGGCGTGTCTCTCATTTTACATCAACAACGATCGATCTTCAGTCCA	120	
DB	175 TGTTCATGGGCGATGTCTGTCTCATTTTATTCAAAGCCACCAACAGCTGGGCTTTCAGTCCA	234	
OY	121 ATGGAATCAGCCAGCTGTGTGCTGGAATAAATACTTCTTTCACCAAACTGATTAT	180	
DB	235 ATGGAGTCTGCAAGTCTGTGCTGGAATAAATAAATTTCTTCTCCACAAAACATGATTAT	294	
OY	161 TTTTATGAAACTACTATTCTGTGTGTGGGTGTGGCCATTTGGGCGACGCTTGTACCTTACA	240	
DB	295 TTTTAAAGCAAACTACATCTGTGTGGGTGTGGCCATTTGGGCGAGACCTTTGACCTTTACA	354	
OY	241 TCTCTCCAAAGCAATGTCTCAACATCCAAAGGATGCGATCTCACAACGAGACCGTACTGTAC	300	
DB	355 TCTCTCCAAAGCAATGTCTCAATATCCAAAGGTGCGATCTCACAACACACGCGCTATTGTAC	414	
OY	301 AACCAATCCCATGACGTTCTGTGATTCATCAATCAACGAGACATCAGTTGGATCTGACAAATTTA	360	

Dd	415	AACAAATCCATGCGGTCCTGATTTCAACCATAGAGACATGAGTGGGATCTGCATTAACCTTA	474		
Qy	361	CCCTGAGCAAGCTAGAGCCACCCCTTCCAGAAATGGATTTGGATGAATTTGGAAATCACCAACT	420		
Dd	475	CCTGAGCAGGCGACGAGCCACCCCTTTCAGAAATGSAATTTGGATGAATTTAGAGTACCCACT	534		
Qy	421	CACAACCCCCAAAAGAGCTGAGCTATTGAGCACTGGTTTTAACCTGACCTGCACTTACCGCGCT	480		
Dd	535	CACACCCCCCAAAGAGTGGCATTTGAACACTTGTTCACCTGACTGCTACTATATGCGCGT	594		
Qy	481	GATTCAGATATATCCAAAGTCCCTTATAGGCTTCTTGACGGTATAGACAAATNCCCTGTGTTT	540		
Dd	595	GATTCAGATATATCCAAAGTCCCTTATAGGCTTCTTGACGGTATAGACAAATNCCCTGTGTTT	654		
Qy	541	GAAGTGCCAACAAGAGAAATTTGGTGGTGGTGTGATGATACGAGAACCCGTAGCAT	600		
Dd	655	GAAGTGCCAACAAGAGAAATTTGGTGGTGGTGTGATGATACGAGAACCCGTAGCAT	714		
Qy	601	GCCAGAGTCAAGATTAATCAATGAGCTAAGCAAAAAGCATTTGAATCCATACCTACGGGCA	660		
Dd	715	GCCAGAGTCAAGATTAATCAATGAGCTAAGCAAAAAGCATTTGAATCCATACCTACGGGCA	774		
Qy	661	GCATTTGGAGAAATATGTCAATGATTAATAATTTGATTCTTACCAATATCTGCTGTAAATTT	720		
Dd	775	GCATTTGGAGAAATATGTCAATGATTAATAATTTGATTCTTACCAATATCTGCTGTAAATTT	834		
Qy	721	TATCTTCCCTTGGAAATTAATCCACCAAGATTAACATACGAAAGGCTATACCAATGCT	780		
Dd	835	TATCTTCCCTTGGAAATTAATCCACCAAGATTAACATACGAAAGGCTATACCAATGCT	894		
Qy	781	TTTCTGGCTGCTCTGTACCTGTTGTGTCTGGACACATCTAGGAAACATATGAGAATTTAT	840		
Dd	895	TTTCTGGCTGCTCTGTACCTGTTGTGTCTGGACACATCTAGGAAACATATGAGAATTTAT	954		
Qy	841	ATTCAGACAGTTCATTCATTCATATGATGGAAGATTAATTACTCTCCAGTACGCTAACAAAG	900		
Dd	955	ATTCAGACAGTTCATTCATTCATTCATATGATGGAAGATTAATTACTCTCCAGTACGCTAACAA	1014		
Qy	901	TATCTGAAGGAAAGTGCAGCAAAAACAAATTAAGTTATTAACCTAGTAACTTAACGTGAGGANG	960		
Dd	1015	TATCTGAAGGAAAGTGCAGCAAAAACAAATTAAGTTATTAACCTAGTAACTTAACGTGAGGANG	1074		
Qy	961	GATTTCACTGTAATTCCTCCACGATTTTGGGATTCACATGACATGATTTGGCTTGGCATCAT	1020		
Dd	1075	GATTTCACTGTAATTCCTCCACGATTTTGGGATTCACATGACATGATTTGGCTTGGCATCAT	1133		
Qy	1021	GTAAAAAGGCATCAAGAAATATAAGTCGTGTGAATTTAGAGAAATGTTTGGAAATTA	1080		
Dd	1135	GTAAAAAGGCATCAAGAAATATAAGTCGTGTGTGAATTTAGAGAAATGTTTGGAAATTA	1194		
RESULT 5					
ID	AAT59505	AAT59505 standard; cDNA; 1814 BP.			
AC	AAT59505;				
XX	06-MAY-1997	(first entry)			
DT					
XX	Murine myeloid-lineage alpha-(1,3)-fucosyltransferase cDNA.				
DE					
XX	Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM;				
KW	septic shock; septicemia; therapy; ss.				
XX					
OS	Mus sp.				
XX					
XX					
FT	Key	Location/Qualifiers			
FT	CDS	325..1353			
XX	/*lag= a				
PN	W09640881-A1.				
DD	19-DEC-1996				

PF 08-MAY-1996; 96WO-US06427.
 XX
 PR 07-JUN-1995; 95US-0483151.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 FI Holgersson J, Seed B;
 XX WPI: 1997-108639/10.
 DR P-PSDB: AAM1820.
 XX
 PT New murine alpha-(1,3)-fucosyltransferase - for fucosylating an
 XX antibody to protect mammals against e.g. septic shock or septicemia
 PS
 XX Claim 6; Fig 3; 58pp; English.
 CC A cDNA clone (AA159505) codes for murine myeloid lineage alpha-
 CC (1,3)-fucosyltransferase (AAW1820), an enzyme which has a relatively
 CC strict substrate requirement for sialylated N-acetylglucosamine,
 CC which can account for the presence of the sialyl-le^x epitope on
 CC murine cells, and which is more effective than Fuc-TIV in support
 CC of E-selectin-mediated COS cell adhesion. It was isolated from a
 CC murine myeloid 32D c13 cell cDNA library by identification of a
 CC clone capable of directing the expression of sialyl-le^x
 CC determinants. Transformed host cells (e.g. 32D c13 or human 293
 CC cells) expressing the alpha-(1,3)-fucosyltransferase, and pref-
 CC erentially fucosylating substrates such as human Fuc-TIV (see also AA159506),
 CC can be used to fucosylate an antibody (e.g. IgG or IgM) for use in
 CC protecting an animal against an adverse immune reaction, esp. septic
 CC shock or septicemia.
 XX
 XQ Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other;
 XX

Query Match	11.8%;	Score 127.4;	DB 18;	Length 1814;
Best Local Similarity	52.1%;	Pred. No. 3e-26;		
Matches 334;	Conservative 0;	Mismatches 301;	Indels 6;	Gaps 2

Oy	269	GAGCAGCATCTCACAAAGGAGACCGTTTCACGTGTATACAAACAAATCCCATGCGATGTTGCATGATCAGC	328
Db	548	GCTGCCTCGTGTAGTGTAAACCGGAGCGCTGCTAGCAGGTGTATGTGTGTGTCTTCACGC	607
Oy	329	ACCCAGAGCATCAGTTGGGATCTGTACAAATTTTACCCTGAGAAAGTATGAGCCACCGCTTCACGA	388
Db	608	ACCGTAGCTGCACAAACCCGGGCATCTCTCTCTACCCCTGGAGCCAGAGGCGACACAGCGACAC	667
Oy	389	AATGGATTTTGGATGAATTTGGATTCACAAACCTACACCTCCCAAAAAGATGGCATTTAGC	448
Db	668	CTTGGGTCTGGGCGCTCCATGTAATCGCCAGTAATACCATCTGCTCCCATCGCTTCGGG	727
Oy	449	ACCTGTTTAAACCTACACCTACCTTTACCGCGCGGATTTGACATATTCACATGCTGTATGCT	508
Db	728	GCATCTTTCACACTGGGTGCTAGCTATGCGGTGTATTCATCATATCTTTGATCCCTACGGTC	787
Oy	509	TCTTGACGGGTAAAGACAAATCCCTGTGTGTTTAAGTCCCAACGACAAAGAAATTTGTTGT	568
Db	788	CGTTTGAGCCCTCTCTCTGGGCG---CACATGCCCATCTACCGGCGCCAAAGACAGATGGCTG	844
Oy	569	GCTGGGTGTGTAGTAACTGCGAACCCCTGAGCATGCGCAGATCAAGTATTTGCAATAGACTAA	628
Db	845	CTGTGGTATATAGCAATTTCTCAGAGGCGGACGACGATGCAAAAGCTGTACCGGACGCTGG	904
Oy	629	GCAAAAGCATGAAATCCATCACTACGGGCAACCATTTGAGATATGTGCAATGTATATA	688
Db	905	CCCCCATCTGCACGTGTGAATGTGTGTGGCTGCGGCACAGCGAGCGCCCTATGCGGCTAATTT	964
Oy	689	ATTGTGATTCCTACCATATCTGTGTGTAAATTTTATCTTTCTTTGAAAATTCATATCCACA	748
Db	965	GCTGTGCTGCCCATCTTTGGCCCGGTACCGCTTCTACCTGCGCTTTTGAGAACTACAGCAATC	1022
Oy	749	AGGATTTACATACGGAAGAACTAT--ACAATGCTTTTGTGGCTGAGCTCTGTACCTGTGTG	805
Db	1025	GGGAGCTACATCTACGAGAAATTTCTGGGCGCATACCCCTGGCGGCTGTGTGTCTGTACCCGATGG	1088

QY	806	TTTGGGACCATCTTGGGAAAAATCTAGAAATATATATTCACAGCAATTCATTCATCATG	865
Db	1085	CCTGTGGACCTCTCTGGGGCCACCTACAGAGCTTTTGTGCACACAGATCCCTTTGTACACG	1144
QY	866	TGGAGATATATTAATCTCTCCAGTAGAGCTAGCAAAATATCTG	906
Db	1145	TGGACACATTCAGCTCTCTCCCGCTAACTGGCTGTCTTCCTG	1185

RESULT 6
AAT85219
ID AAT85219 standard; DNA; 3594 BP

AC AAT85219;

DT 02-MAR-1998 (first entry)

Mouse alpha-fucosyltransferase Fuc-TVII gene.

KW Alpha-fucosyltransferase; Fuc-TVII gene; mouse

KW inhibitor; ds.

Mus musculus.

FH	Key	Location/Qualifiers
FM	005	0000

```
/*tag= a
```

FT exon 465..477

```

478 /number= 1
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399

```

```
/*tag= c
```

FT	exon	670..1149
FT	exon	670..1149

```

FT      /number= 2      995
FT      /codon_start= 995 998

```

ET 1150.:1946

FT	exon	/number = 2	1947	1959
FT				

```
ET      / *tag= 1
ET      /number= 3
```

```

ET /coach_start= 1947..1950
ET /note= "alternative A3"

```

1960.:2063

FT	number = 3
FT	2067 - 3552
exon	

```

E1 /cag= 1
E1 /number= 4
E1

```

```

/cool_start= 2120...21
/notes="alternative A"

```

AA
AA
PN W09732889-A1.

PD 12-SEP-1997.

07-MAR-1997; 97WO-US03184.

08-MAR-1996; 96US-0613098.

PA (UNMI) UNIV MICHIGAN.

PI Gersten KM, Lowe JB, Natsuka S;

DR WPI; 1997-470535/43.

XX

PT Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g.

PT screening for inhibitors used as antiinflammatory agents
 XX
 PS Claim 1; Fig 2; 91pp; English.
 CC This DNA sequence comprises the mouse Fuc-TVIII gene encoding
 CC alpha-fucosyltransferase. It was isolated from a 373 kb genomic
 CC library by screening with a probe from the catalytic domain of the
 CC human Lewis alpha(1,3/1,4)fucosyltransferase (Fuc-TVIII).
 CC Examination of the mouse Fuc-TVIII locus identified 3 Met codons
 CC that may initiate translation of alpha(1,3) fucosyltransferases
 CC with different cytosolic domains encoded by exons 2 and/or 3, but
 CC with identical Golgi-localised catalytic domains encoded by exon 4
 CC (see AM27138 and AM26671). The proteins and nucleic acids can
 CC be used to construct animal cell lines able to post-translationally
 CC modify oligosaccharides on cell surface, intracellular or secreted
 CC proteins and lipids, to isolate reagents for efficient enzymatic
 CC production of oligosaccharides, to generate antibodies to
 CC glycosyltransferases useful as diagnostic reagents, to screen for
 CC fucosyltransferase inhibitors and inactivators, especially those
 CC that act as antiinflammatory agents, for genotyping individuals at
 CC the fucosyltransferase locus and for in vitro synthesis of sialyl
 CC Lewis X tetrasaccharide.
 XX

Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other;

Query Match 11.8%; Score 127; DB 18; Length 3594;
 Best Local Similarity 52.3%; Pred. No. 5.2e-26;
 Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

QY 269 GATGCCATCTCACACGACCGCTTACCTACACAAATCCATGAGTTTCATTCATC 328
 DB 2277 GCTGCCGTCGTAGTGTACACGAGCGCTCTAGCCAGTCTGTGCTTCCTACCC 2336
 QY 329 ACCGACATCATGTTGGATGTGACAAATTTACCTCAGCAGTCCACCTCCAGA 388
 DB 2337 ACCGTAGCTGCACAAACCCGCGCATCTCTCTACCCCTGACACAGAGCGGACAGC 2396
 QY 389 AATGATTTGATGAAATTTGGAATCACAATCACAATCCCAAAAGATGGCATTTGAGC 448
 DB 2397 CTTGGGCTTGGGCTTCATGAAATGCCACAGTAAATCCATGATGTTCCATCGCTCCGG 2456
 QY 449 ACTTGTTAACCTGACTCTGACTACCGCGCTGATTCATATATCCAAAGCTTATGCT 508
 DB 2457 GATCTTCACTGAGGTGCTGAGTATCGGCTGATTTCAATATCTTTGACCTTACGGTTC 2516
 QY 509 TCTTACGCTTAAGCACAAATCCCTTGTGTTGAAGTGCACAAAGAAATTTGTTG 568
 DB 2517 GCTTGAGGCTCTCTCTGGGCC---CACATCCCACTACCGGCAAAAGCAGATGGCTG 2573
 QY 569 GCTGGTTGTGATACCTGAAACCTGAGCAATGCGAGTCAAGTATTCATATGACTAA 628
 DB 2574 CCTGGGTGATCAGCAATTTCCAGAGCGGACGAGCGTGAAGTGTACCGGCACTGG 2633
 QY 629 GCAAAAGCATTTGAATTCATACCTAGCGCAAGCATTTGGAGCATATGTTCATGATAAAA 688
 DB 2634 CCCCTCATCTGAGGTGATGTTGTCGCGCGACGCGCGCCCATGCGCTAATTT 2693
 QY 689 ATTGATTCCTACATATGCTGTGTAATTTTATCTTCTTGAATTTCAATCCACA 748
 DB 2694 GTCTGCTGCCCATTTGGCGCGGTACCGCTTCTACTGGCTTTTGAAGACTCACACATC 2753
 QY 749 AGGATACATCAACGAAAGCTAT--ACAAATGCTTTTCTGGCTGCTGTACCTGTTG 805
 DB 2754 GGGACTACATCACTGAGAGTCTGGCGCAATGCCCTGGCGGTGTGTACCCGGTGG 2813
 QY 806 TTCTGGACCATCTAGGAAACTATGACAATTTATTTCCAGCAGATTTATTCATTCATG 865
 DB 2814 CCGTGGGACCTCTCGGGCCACCTACAGAGCTTTGTGCACACAGATGCTTTGTACAG 2873
 QY 866 TGGAGATTTAATCTCTCCAGTAGAGCTAGC 896
 DB 2874 TGGACGACTTACGCTCTGCCCGTGAAGCTGGC 2904

RESULT 7
 AAQ77732
 ID AAQ77732 standard; cDNA to mRNA; 1701 BP.
 XX
 AC AAQ77732;
 XX
 DT 22-JUN-1995 (first entry)
 XX
 DE Human alpha-1,3-fucosyltransferase cDNA.
 XX
 KW Alpha-1,3-fucosyltransferase; sialylated Lewis blood group antigen;
 KW Lewis x; Lewis a; glycosylation; glycoprotein; glycolipid; ds.
 XX
 OS Homo sapiens (THP-1 cells).
 XX
 FH Key Location/Qualifiers
 FT CDS 123..1151
 FT /tag= a
 FT /product= alpha-1,3-fucosyltransferase
 PN W09423021-A.
 XX
 PD 13-OCT-1994.
 XX
 PF 28-MAR-1994; 94WO-JP00496.
 XX
 PR 29-MAR-1993; 93JP-0069016.
 XX
 PA (KYO) KYOMA HAKKO KOGYO KK.
 PI Hanai N, Kurata K, Nishi T, Sasaki K;
 XX
 DR MPI: 1994-333183/41.
 XX
 DR F-PSDB; AAR63215.
 PT Alpha-1,3-Fucosyl-transferase DNA and protein - useful for prodn.
 PT of active sugar chains, e.g. sialylated Lewis X
 PS Claim 3; Page 99-102; 155pp; Japanese.
 XX

Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;

Query Match 10.6%; Score 114.4; DB 15; Length 1701;
 Best Local Similarity 50.8%; Pred. No. 1.6e-22;
 Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

QY 256 TTCACATCCAAAGATGCCATCTCACACGAGCGTTCTACTGACAAATCCATGCA 315
 DB 333 TACGGATGCGCCGCTGCGACCTGATGCCAATCGAAGCTGTGCGCAGCGCGACGCC 392
 QY 316 GTTCTGATCCATCCAGGAGCATCACTTGGGATCTGACAAATTTACTCAGCAAGCTAG 375
 DB 393 GTGCTTTCACACCGCGAGCTGCGAGACCGGCGGTCCACCTGCGCCGCGACCGG 452
 QY 376 CGACCTTCAGAAATGATTTGATGAATTTGGAATTCACCACTCACTCCCAAAAG 435
 DB 453 CCGGAGGAGGAGCGCGGTGCGGCTGCATGAGATCTTCATGACACACCGAGCGCTC 512
 QY 436 AGTGGCATTTGAGCATTTGTTAACTGACTGTGACTTACCGCGCGATTCAGATATCAA 495
 DB 513 AGCCACCTCCAGAGCATTTTCAACTGGGTGAGCTGACCTGCGCGCATCTGCTTT 572
 QY 496 GTGCTTATGGCTTCTTGAAGGTAAAGCACAATCCCTTGTGTTGAATGCGCAAGCAA 555
 DB 573 GTGCTTATGGCGCGCTGGAAGCCCACTGGGGGCGCT---CGCACCGCTGCGACCAAG 629

QY	556	GAAAAATGTGTGTGGTGTGTGATTAATGCAACCCGACATGCCAGAGTAAAT	615
Db	630	ACAGAGGTGGCCGCTGGGTGGTACCAACTTCCAGAGCGCACTGGCTGCCAGGCTG	689
QY	616	TAAATATGACCTPAAACCAAAAGCATTTGAATTCATACCTACGGCGCAAGCATTTGGAGAAATAT	675
Db	690	TACCGGACAGTGGCGCCCTCAATCTGGGTGGATGTCTTTGGCGGCGCCAAAGAGCGCA	749
QY	676	GTCATGATGATAAAAATTTGATATCTTACCATATCTGCTTTGTAATTTTATCTTTCCPTTGAA	735
Db	750	CTGTGGCCGACGTGCTGTGTCTCCACCGTGGCCACGATACCGCTCTTACCTCTTCCPTTGG	809
QY	736	AATTCATCCACAGAGATTTACATCAGGAAAAAGCTAT--ACATATGCTTTTCTGGCTGGC	792
Db	810	AACCTTCACACACCGGACGTACATTTACGAGAAATTTCTGGCGCCACAGCACTGTGGCTGGC	869
QY	793	TCTGTACCTGTTGTTCTGGGACCATCTAGGAGAAACCTATGACAAATTTATATTCACAGCAT	852
Db	870	ACTGTCCAGTGTGCTGTGGGCCCCACGGGGCCACCTATGAGAGCCTTGTCGCCGGCTGAC	929
QY	853	TCAATTCATTCATGTGAGAGATTATACCTCCACGAGACTAGC	896
Db	930	GCCCTGCGATGTGATGATGACTTTGGCTGACGCCGAGAGACTGGC	973

RESULT 8
AAF21136
ID AAF21136 standard; DNA; 1701 BP.

DT 14-MAR-2001 (first entry)

Human low adenosine antisense oligonucleotide related sequence #2703.

KM low adenoma antisease oligonucleotide; phosphorothioate; allergy;
 human; always disorder; bronchocarcinosis; lung inflammation;
 KM sufficient depletion; respiratory; bronchodilator; antiinflammatory;
 immunosuppressive; antihistamic; analgesic; hypotensive; cytostatic;
 respiratory obstruction; pulmonary obstruction; impeded respiration;
 KM sufficient hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis
 KM pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KM chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KM cancer; ss.

OS Homo sapiens.

PN WO200062736-A2.

PD 26-OCT-2000.

PF 24-MAR-2000; 2000WO-US08020.

PR 06-APR-1999; 99US-0127958.

PA (UYEC-) UNIV EAST CAROLINA.

XX

XX

DR WPI; 2000-679539/66.

PT Low adenosine (A) content antisense oligonucleotides which do not

PT cancers and respiratory obstructions -

PS Disclosure; Page 1034; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,

immunosuppressive, antilasthmatic, hypotensive and cyostatic activities. The antileukemic oligonucleotides and (1) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antileukemic oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergies) and/or obstructive hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impaired respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAr18434 to AAr21543 represent human polynucleotide fragments and antileukemic oligonucleotides used in the exemplification of the present invention.

Sequence 1701 BP; 290 A; 527 C; 589 G; 295 T; 0 other;

Query Match	10.68;	Score 114.4;	DB 21;	Length 1701;
-------------	--------	--------------	--------	--------------

Matches 327; Conservative 0; Mismatches 311

Matches	327;	Conservative	0;	Mismatches	311;	Indels	6;	Gaps	2;
---------	------	--------------	----	------------	------	--------	----	------	----

256 TTCACATCCAAGGATGCCATCTCACACGAGCCGTTCACTGTACACAATAATCCCATGCA 315

Db 333 TACGGCATCGCCCGCTGCCACCTGAGTGCCAACCGAAGCCTGCTGGCCAGCGCGGACGCC 392

QY 316 GTTCTGATCCATCACCAGACATCAGTTGGGATCTGACAAATTACCTCAGCAAGCTAGG 375

Db 393 GTGGTCTTCCACCACCGGAGCTGCAGACCCGGCGTCCACCTGCCCTGGCCAGCG 452

376 CCACCCCTTCAGAAATGGATTGGATGAATTTGGAATCACCACCTCACACTCCCAAAG 435

Db 453 CCGGAGGGCAGCCCTGGGTGTGGCCCTCCATGGAGTCTCCTAGCCACACACCACGGCCTC 512

436 AGTGGCATTGAGCACTTGTTAACCTGACTCTGACTTACCGCGGTGATTCAGATATCCA 495

Db 513 AGCCACCTCCGAGGCATCTTCAACTGGGTGCTGAGCTACCGGGCGGACTCGGACATCTT 572

496 GTGCTTATGGCTTCTTGACGGTAAGCACAATCCCTTCGTGTTGAAGTGCCACGCAA 555

Db 573 GTGCCCTATGCCCGCCTGGAGCCCACTGGGGCCCT--CGCCACCGCTGCCAGCCAAG 629

556 GAGAAATTGGTGTCTGGTTGTGAGTAACTGGAACCTGAGCATGCCAGAGTCAAGTAT 615

Db 630 AGCAGGGTGGCCGCTGGGTGTCAGCACTTCAGGAGCGGCACTGCCGTGCCAGGCTG 689

616 TACAATGAGCTAAGCAAAGCATTTGAATTCATACCTACGGGCAAGCATTGGAGAATAT 675

Db 690 TACCGGACGCTGGCGCCTCATCTGCGGGTGATGTTTGGCCGTGCCAATGGACGGCA 749

676 GTCAATGATAAAATTGATTCCTACCATACTCTGTAATTTTATCTTCCCTTGA 735

Db 750 CTGTGCGCCAGCTGCCTGGTGCACACCGTGGCCAGTACCGCTTCTACCTGTCTTGAG 809

736 AATTCAATCCACAAGGATTACATCAGGAAAGCTAT--ACAATGCTTTTCTGGCTGGC 792

Db 810 AACTCTCAGCACC GCGACTACATTACGGAGAATTCTGGCGCAACGCCACTGGTGCTGGC 869

793 TCTGTACCTGTTGTTCTGGGACCATCTAGGGAACCTATGAGAATTATATTCACGACGAT 852

D6 870 ACTGTGCCAGTGGTGGGGCCCCCACCAGGGGCCACCTATGAGGCCCTTCGTTGCCGGCTGAC 929

853 TCATTCATTCATGTGGAAGATTATACTCTCCAGTGAGCTAGC 896

Db 930 GCCTTCGTCATGTGATGACTTTGGCTCAGCCCGAGAGCTGGC 973

RESULT 9

AAA35014
ID AAA35014 standard; DNA; 1701 BP.

AC AAA35014;

DT 28-JUL-2000 (first entry)

DE Human adenosine receptor related polynucleotide SEQ ID NO: 2703.

Human; adenosine receptor; low adenosine antisense oligonucleotide;
phosphorothioate; impaired respiration; inflammation; allergy;
allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
respiratory distress syndrome; pain; cystic fibrosis; emphysema;
pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

OS Homo sapiens.

PN WO200009525-A2.

PD 24-FEB-2000.

PE 03-AUG-1999; 99WO-US17712.

PR 03-AUG-1998; 98US-0095212.

PA (UYEC-) UNIV EAST CAROLINA.

PI Nyce JW;

DR WPI: 2000-205971/18.

New antisense oligonucleotides useful for treating e.g. pulmonary
vasoconstriction, inflammation, allergies, asthma, hypertension,
bronchitis, emphysema, respiratory distress syndrome, ischemia or
cancers -

Disclosure: Page 957-958; 1343pp; English.

The present invention describes a new composition comprising an
antisense oligonucleotide (ON) with low adenosine (up to 15%), which
targets nucleic acids involved in bronchoconstriction, allergies, and/or
inflammation. The ON can have antiinflammatory, antiallergic,
antiasthmatic, cytostatic and analgesic activities. The compositions are
useful for the treatment of diseases associated with inflammation,
impaired airways, including lung disease and diseases whose secondary
effects afflict the lungs of a subject. They can be used for treating
e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
asthma, impaired respiration, respiratory distress syndrome, pain, cystic
fibrosis, pulmonary hypertension, emphysema, chronic obstructive
pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
carcinomas, and cancers which may metastasise to the lungs, including
breast and prostate cancer. The reduction of the adenosine content of
the ONs reduces side effects. The A-containing ONs break down with the
release of deoxyadenosine which activates adenosine receptors causing
bronchoconstriction and inflammation. AAA3213 to AAA5312 represent the
nucleotide sequences given in the sequence listing from the present
invention, which correspond to SEQ ID NO:1 to 2815, and then the last
185 sequences are also called SEQ ID NO:1 to 185, but the sequences
differ from the previously named sequences. SEQ ID NO:11 to 1680
(AAA32323 to AAA33992) are specifically claimed ONs from the present
invention. N.B. Sequences given in the disclosure of the present
invention do not match up with their corresponding SEQ ID NO. sequences
given in the sequence listing.

Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;

Query Match 10.6%; Score 114.4; DB 21; Length 1701;
Best Local Similarity 50.8%; Pred. No. 1.6e-22;
Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

QY 256 TTCACATCCAMAGATGCCATCTCACACGACCGCTTCACTGTACACAAATCCATGCA 315
DB 333 TACGGATCGCCCGCTGCCACCTGATGTCACACCGAAGCTGTGGCCAGCCGAC 392
QY 316 GTTCGATCCATCACCAGACATCATCTTGGGATCTGACAAATTTACTCAGACGTAG 375
DB 393 GTGGTTTCCACACCGACGATGCTGACAGACCAGGCTCCACCTGCTCCGCGCAGGG 452
QY 376 CCACCTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
DB 453 CCGGAGGGGACCGCTGGGTGTGGCTCCATGATGATCTCTTACACACACCGACGCTTC 512
QY 436 AGTGCATTGACGACCTGTTTAACTGACTGCTGACTGCTGACTGCTGCTGCTGCTGCT 495
DB 513 AGCCACCTCCGAGGACATCTTCAACTGGTGTGAGCTGAGCTGAGCTGAGCTGAGCTG 572
QY 496 GTGCCCTTATGCGTCTGTAAGCGTAAGCAAAATCCCTGCTGTTTAACTGCGCAACAA 555
DB 573 GTGCCCTATGCGCGCTGAGAGCCCACTGGGGCCT---CGCACCGCTGCGCAAG 629
QY 556 GAGAAATTTGTTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 615
DB 630 ACAGGCTGGCGCGCTGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 689
QY 616 TACATGACCTAAGCAAAAGACATGTAATCATCTGACGCGGACGATGAGATAT 675
DB 690 TACCGGACACTGGCGCTCATCTGCGGGTGTGATGATGATGATGATGATGATGATGATG 749
QY 676 GTCAATGATAAATAATTTGATTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
DB 750 CTGTGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809
QY 736 AATTCATTCACAGATGATTCATCAACGAAAGCTAT---ACAAAGCTTTTGTGCTGCG 792
DB 810 AACTGCACACACCGCGACATCATGAGGAAATTTGTGGCGACACCATGCTGCTGCTGCG 869
QY 793 TCTGTACTGTTGTTCTGGGACCATCTAGGAAACTATGAGAAATTAATTCACAGCAT 852
DB 870 ACTGTGCCAGTGTGTGGGCGCCCGCACGCGCACCTATGAGAGCCCTTCTGCGGCTGAC 929
QY 853 TCATTCATTCATGATGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 896
DB 930 GCCTTCGTCATGTGATGACTTTGGCTCAGCCCGAGAGCTGGC 973

RESULT 10
AAE21137
ID AAE21137 standard; DNA; 6944 BP.

AC AAE21137;

DT 14-MAR-2001 (first entry)

DE Human low adenosine antisense oligonucleotide related sequence #2704.

Human; low adenosine antisense oligonucleotide; phosphorothioate; allergy;
human; airway disorder; bronchoconstriction; lung inflammation;
surfactant depletion; respiratory; bronchodilation; antiinflammatory;
immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
respiratory obstruction; pulmonary obstruction; impaired respiration;
surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
pulmonary hypertension; emphysema; pulmonary transplantation rejection;
chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
cancer; ss.

OS Homo sapiens.

CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antisthmatic, cytosolic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, leukemias, lymphomas,
 CC pulmonary disease (COPD), and cancers which may metastasise to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing the
 CC bronchoconstriction and inflammation. AAA32313 to AAA3312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32323 to AAA3392) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.

XX Sequence 6944 BP; 1218 A; 2193 C; 2223 G; 1310 T; 0 other;

Query Match 10.6%; Score 114.4; DB 21; Length 6944;

Best Local Similarity 50.8%; Pred. No. 3e-22;

Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

QY 256 TTCAATCCAAAGATGCCATCTCACACGACCGCTTCACTGACAAACAATCCATGCA 315
 DB 5576 TACGGATGCGCCGCTGACCACTGATGCCAACGACCTGCTGGCCAGCCGCGCC 5635
 QY 316 GTCTGATCCATCCAGAGACATCATGTTGGATCTGACAAATTTACTCTGCAAGTAGG 375
 DB 5636 GTGCTCTTCCACACCGGACGCTGACAGCCCGGCTCCACTGCCCTGGCCAGCGG 5695
 QY 376 CCACCTTCCAGAAATGGATTTGGATGAAATTCACCACTGACACTCCCAAAAG 435
 DB 5696 CCGGAGGAGGACGCGCTGGGCTGCGCTCCATGAGTCTCTAGCCACACCGGCGCTC 5755
 QY 436 AGTGGCATTTAGACATCTTTTAACCTGACTCTTACCGCGGTGATTCAGATTCACA 495
 DB 5756 AGCCACTCTCGAGGACATCTTCAACTGGTGTCTAGCTACCGGCGCATCTGACATCTT 5815
 QY 496 GTGCTTATGAGGCTTTGAGCGGTAAGCAACAATCCCTGCTGTTGAAGGCCAAGCAA 555
 DB 5816 GTGCTTATGAGGCTTTGAGCGGTAAGCAACAATCCCTGCTGTTGAAGGCCAAGCAA 5572
 QY 556 GAGAAATGCTGTGCTGCTGTGATTAACCTGAGACCTGAGCATCCAGATCAAGTAT 615
 DB 5873 AGCAGGCTGCGCCGCTGGGCTGATCAACCTTCCAGAGGCGGAGAGTGGCTGCAAGCTG 5932
 QY 616 TACAAATGAGCTAAGCAAAAAGCATTTGAATCACTACGCGGCAAGCATTTGAGAAATAT 675
 DB 5933 TACCGGACGCTGCGCTCATCTGCGGGTGAGATGCTTTGGCCGCTCAATGAGACGGCCA 5992
 QY 676 GTCAATGATTAATAATTTGATTTCTTACCATATCGTTGTAATTTTATCTTCTTTGAA 735
 DB 5993 CTGTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6052
 QY 736 AATTCAATCCACAAGATTAACATCAGGAAAGCTAT--ACAATGCTTTTCTGCTGCG 792
 DB 6053 AACTCTACAGACCGGACGATCACTTACGAGAAATTTCTGGGCAAGGACATGAGGCTGCG 6112
 QY 793 TCTGTACCTGTTTCTTGGGACCATCTAGGAAAACATATGAGAAATATATTCACGACAT 852
 DB 6113 ACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6172
 QY 853 TCATTCATTCATGATGAGAAATTAATTAATCTCCAGTAGAGTAGC 896
 DB 6173 GCCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6216

RESULT 12
 AA056905
 ID AA056905 standard; DNA; 1654 BP.

XX AA056905;

DT 26-JUL-1994 (first entry)

DE pCDNA1-alpha-(1-3)-Fuc-TV1 genomic DNA.

KW Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;

KW surface; oligosaccharide; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 128..1208

FT /tag= a

PN W09402616-A.

PD 03-FEB-1994.

XX 20-JUL-1993; 93MO-US06703.

PR 20-JUL-1992; 92US-0914281.

PA (UNMT) UNIV MICHIGAN.

XX Lowe JB;

PI WPI; 1994-048874/06.

DR P-PSDB; AAR45933.

XX DNA fragment encoding a glycosyltransferase - can be used for in

PT vitro reactions to modify cell surface oligosaccharide(s) e.g.

PT blood gp. determinants, to protect against transplant rejection

PS Claim 3; Fig 7; 249pp; English.

XX The sequence is that encoding human GDP-Fuc: [beta-D-Gal (1-4)]-D-

CC GlucNAc alpha (1-3)-fucosyltransferase. The enzyme produced by the

CC DNA may be non glycosylated. This prevents premature loss of enzyme

CC activity. It can also be used in in vitro reactions to modify cell

CC surface oligosaccharide mols. e.g. blood group determinants.

XX See also AA056906-12.

XX Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match 10.5%; Score 113.4; DB 15; Length 1654;

Best Local Similarity 51.2%; Pred. No. 3e-22;

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

QY 196 ATCTGTGTGTTGGTGGCCATTTTGGGACAGACCTTTGACCTTACATCTGCGCAAGCAATG 255
 DB 318 ATCTGTGTGTTGGTGGCCATTTTGGGACAGACCTTTGACCTTACATCTGCGCGCTGCTGAGATG 377
 QY 256 TT--CAACATCCAAAGATGCTTCATCAGACGACCGCTTACTGTACAAACAATCCAT 312
 DB 378 GTGCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
 QY 313 GCAGTTCTGATCCATCAGCAGACATCAGTTGGATCTGA---CAATTTACTCTGACGAA 369
 DB 438 GCGGTATGCTGTCACACCGAGAGCTCATGTACACCCCAAGTCCAGCTCCACGCTCC 497
 QY 370 GCTAGGCAACCTTCCAGAAATGGATTTGGATGAAATTTGGAAATCACCACACTGCTCC 429
 DB 498 CCGAGGCGGAGGAGGAGGATGATCTGTTGATGATGATGATGATGATGATGATGATGATG 557
 QY 430 CAAAAGATGAGCTTACGACATTTTAACTGATCTGACATTTACCGCGGTGATTCAGAT 489

```
Db 558 CAGCTGAAGCATGAGGATGACTCATCTCACCATGCTTCTACCGCAGCATCTCCG 617
OY 490 ATCCAGAGGCTTATGAGCT-----TCTTGACGGTAGACCAAAATCCCTGCTGTTT 540
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 618 ATCTCAGCGCTTACGGTGGCTGGAGCGCTGGTCCGCCAGCGCTCCACCCACCGCTC 677
OY 541 GAAGTCCCAAGCAAGAAATGTTGTGCTGGGTTGTGTAAGTACCTGAAACCTTACAT 600
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 678 AACCTCTGGCCCAACACCGAGCTGTGGCTGGCACTGGGGGCCAAATCTCC 737
OY 601 GCCAGAGTCAAGTATTACATGAGCTTAAGCAAAAGCAATTTGATACCTACCGGCAA 660
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 738 GCCAGGGTGGCTTACTACAGAGCTGAGGCCATCTCAAGGTGAGCTGACGG 794
OY 661 GCATTTGGAGAAATATGATGATTAATAAAATTTGATTCCTCAATCTGCTTGAATTT 720
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 795 CGCTCCCAAGACCCCTGCCCCAGGAAACATGATGAGAGCGCTGCTCCGTAACAAGTTC 854
OY 721 TATCTTCCCTTGAATAATTCATCCACAGGATTCATCCAGGAAAGCTAT---ACAAT 777
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 855 TATCTGCGCTTGAGAACTCCCTGACCCGACTACATCCAGAGAACTGTGAGAGAAC 914
OY 778 GCTTTCTGGGCTGTGTGTAAGCTGTTGTTCTGAGAACATCTGAGAAACATGAGAAAT 837
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 915 GCGCTGAGAGCTGTGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
OY 838 TATATCCAGACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 897
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 975 TTCTGCGACCGGACCGCTTCTATCCAGTGGACGACTTCCAGAGCCCAAGGACCTGGCC 1034
OY 898 AAGTATCTGAGAGAGTGCAGCAAAACATTAAGTTATACCTTAAGTTAACTGAGAG 957
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1035 CGGTACCTGCAGAGAGCTGCAGCAAGACCGCCGCTACCTGAGTACTTGTGCTGGGG 1094
OY 958 AAG 960
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1095 GAG 1097
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 13
AAT61680
ID AAT61680 standard; DNA: 1654 BP.
XX
AC AAT61680;
XX
DE 19-JUN-1997 (first entry)
XX
KW Human alpha(1,3)-fucosyltransferase (Fuc-TVI) DNA.
XX
KW Alpha(1,3/1,4)-fucosyltransferase; Lewis enzyme; Fuc-TVI;
XX glycosylation; oligosaccharide; blood group; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 129..1208
FT /tag= a
XX
PN WO9709421-A1.
XX
PD 13-MAR-1997.
XX
PF 06-SEP-1996; 96MO-US13816.
XX
PR 08-SEP-1995; 95US-0525058.
XX
PA (UNMT ) UNIV MICHIGAN.
XX
PI Legault DJ, Lowe JB;
XX
DR WPI: 1997-192897/17.
DR P-PSDB: AAW13643.
XX
```

```
PT New recombinant fucosyltransferase proteins - useful for modifying
PT cell surface oligosaccharide structures
PS Example 6; Page 291-292; 329pp; English.
XX
XX A DNA clone (AAT61680) codes for an enzyme (AAW13643) that functions as
CC a GDP-Fuc:(beta-D-gal(1,4)-D-glucNAc alpha(1,3)-fucosyltransferase
CC (Fuc-TVI) which can utilize neutral type II oligosaccharide
CC precursors but not type I glycoconjugates. It was identified in a
CC human lambda phage genomic DNA library using a probe obt. from
CC Lewis enzyme cDNA (see also AAT61675). The Fuc-TVI DNA can be used
CC to construct animal cell lines with specific capabilities with
CC respect to post-translational modification of the oligosaccharides
CC of expressed proteins or lipids, to produce recombinant Fuc-TVI for
CC use in oligosaccharide prodn., to genotype individuals at this
CC fucosyltransferase locus, and in the prodn. of chimelic
CC fucosyltransferases (see also AAW14512-32).
XX
SQ Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;
Query Match 10.5%; Score 113.4; DB 18; Length 1654;
Best Local Similarity 51.2%; Pred No. 36-22;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
OY 136 ATTCGTGTGTGGGTGTGGCCATTGGGCGAGACCTTTGACTTACATCTCGCAAGCAATG 255
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 318 ATCTCTGTGTGAGGTGTGGCTTTTACAAACCATATAGCTGTGCCCTGCTGAGAGATG 377
OY 256 TT---CAACATCCAGAGATGCATCTCACAGAGCGCTTGCATCTGATCAACAAATCCAT 312
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 378 GTGCGTGCAGAGGCTGTGATCAATCACTGCGACCCAGAGGTGTATCCACAGGAGAGC 437
OY 313 GCAGTTCGATCATCCACGAGACATCAAGTGGATCTGA---CAAAATTACCTCAGCAAA 369
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 438 CGGTCATCGTGCACACACGAGAGCTATGTACAACCCAGTCCGACGCTCCACGCTCC 497
OY 370 GCTAGCGACCCCTTCAGAAATGATTTGATGATGATTTGAAATCCAGACGACACTGCC 429
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 498 CCGAGGGGGGAGGGGGGAGCGATGATCTGTTCACAGATGAGTCCCAAGCCGACTGTGG 557
OY 430 CAAGAAGTGGCAATGGACACTGTTTAACTGATCTGATCTGATCTGATCTGATCTGATCTG 489
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 558 CAGCTGAAGAGCATGGAGATCTCAATCTCACCATGTCCTTACGAGAGCTCGAGC 617
OY 490 ATCCAGTGCCTTATGAGCT-----TCTTGACGGTAAAGCAAAATCCCTTGCTGT 540
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 618 ATCTTCAGGCCCTTACGGCTGTGGAGCCGTGGCCGACGCTGCCACCCACCGCTC 677
OY 541 GAAGTCCCAAGCAAGAAATTTGTTGCTGGGTTGTGATTAATCTGGAACCTGAGCAT 600
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 678 AACCTCTCGGCCAAGACCGAGCTGTGGCTGTGGCAGTGTCCAACTGGGGGCCAAATCTCC 737
OY 601 GCCAGAGTCAAGTATTACATGAGCTTAAGCAAAACATTTGAATTCATTAACCTGAGGCAA 660
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 738 GCCAGAGGTGGCTTACTACAGAGCTGAGGCCATCTCAAGGTGAGAGCTGTAGGG---A 794
OY 661 GCATTTGGAGAAATATGATGATTAATAAAATTTGATTCCTACATATCTGCTGTAATTT 720
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 795 CGCTCCCAAGACCCCTGCCCCAGGAAACCATGATGAGAGAGCTGTCCCGGTCAAGTTTC 854
OY 721 TATCTTCTCTTGAATAATTCATCCACAGATTCATTCATTCATTCATTCATTCATTCAT 777
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 855 TATCTGCGCTTGAGAACTCTTTCAGCCCGGATCTACATCCAGGAAGCTGTGGAGAGAT 914
OY 778 GCTTTCTGGGCTGTGTGATCTGTTGTTCTGGAGACCATCTAGGAAAAATCTATGAGAAAT 837
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 915 GCGCTGAGAGGCTGTGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
OY 838 TATATCCAGACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 897
XX 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 975 TTCTGCGACCGGACCGCTTCACTCAGGTGAGAGCTTCCAGAGCCCAAGAGACCTGGCC 1034
OY 898 AAGTATCTGAGAGAGTGCAGCAAAACATTAAGTTATACCTTAAGTTAACTGAGAG 957
```

```

Db 1035 CGGTACCTGCAGAGCTGGACAAAGACACCGCCGCTACTGAGCTACTTTCGCTGGCG 1094
QY 958 AAG 960
    ||
Db 1095 GAG 1097

RESULT 14
AAQ56912
ID AAQ56912 standard; DNA; 1086 BP.
XX
AC AAQ56912;
XX
DT 26-JUL-1994 (first entry)
XX
DE Lewis blood group fucosyltransferase DNA.
XX
KM Glycosyltransferase; fucosyltransferase; GDP-Fuc; In vitro; cell;
KW surface; oligosaccharide; ss.
XX
PN W09402616-A.
XX
PD 03-FEB-1994.
XX
PF 20-JUL-1993; 93WO-US06703.
XX
PR 20-JUL-1992; 92US-0914281.
XX
PA (UNM1 ) UNIV MICHIGAN.
XX
PI Lowe JB;
XX
DR WPI; 1994-048874/06.
XX
PT DNA fragment encoding a glycosyltransferase - can be used for in
PT vitro reactions to modify cell surface oligosaccharide(s) e.g.
PT blood gp. determinants, to protect against transplant rejection
XX
PS Disclosure; Fig 6; 249pp; English.
XX
CC The sequence is that encoding a Lewis blood group fucosyl
CC transferase (Fuc-TIII). The enzyme produced by the DNA may be
CC non glycosylated. This prevents premature loss of enzyme activity.
CC It can also be used in in vitro reactions to modify cell surface
CC oligosaccharide mols. e.g. blood group determinants.
CC See also AAQ56905-11.
XX
SQ Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;

Query Match 10.4%; Score 112.2; DB 15; Length 1086;
Best local Similarity 51.6%; Pred. No. 5.5e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

```

QY 559 AAATTGGTGTCTGGGTTGTGACTGAACCTGAGACCTCCAGAGTCAAGTATTAC 618
    ||||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 571 GAGCTGGTGGCTGGGGGCTGCCAACTGGAAGCCGAGCTACAGGCGCTACTAC 630
QY 619 AATGAGCTAAGCAAAAGCATTTGAATTCATACCTAGGGGCAACATTTGGAAATATGTC 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 631 CAGAGCTGAGGCTCATCTCAAGGTGAGAGTGGAGAGCTCCCAAGCCCTGCCCC 690
QY 679 AATGATATAAATTTGATCTTACCATATCTGCTGTAATTTTATCTTCTTGAATAAT 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 AAGGGAGCCA---TGATGGAGAGCTGTCCCGGTACAACTTCTGAGCTTGAGAAC 747
QY 739 TCAATCCACAAGGATTACATCAAGAAAGCTAT---ACAATGCTTTTCTGGCTGCT 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 748 TCCCTGGACCCCGACTCATACCGAGAACGCTGGAGAGAACGCCCTGGAGGCC 807
QY 796 GTACCTGTTTCTGGGACCATCTAGGAAAACTATGATATATTCACAGCATTTCA 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 GTCCCGGTGGTGGTGGCCCGCCAGCAGAACGATACGAGAGTTCCTGCCACCCGAGGCC 867
QY 856 TTCATTCATGTGGAGATTAATTAATCTCCAGTGAAGCTACCAAGTATCTGAAGAAATC 915
    ||||| ||||| ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 868 TTCATTCACGTGGAGATCACTTCCAGAGCCCAAGAGACTGGCCCGGTACTGCAAGACTG 927
QY 916 GACAAAAACAATAAGTATATACCTTAATTAATTAATTAATTAATTAATTAATTAAT 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 GACAAAGACCAACGCCGCTACCTGAGCTACTTCCGTGGCGGGAG 972

RESULT 15
AAT80111
ID AAT80111 standard; cDNA; 2042 BP.
XX
AC AAT80111;
XX
DT 01-DEC-1997 (first entry)
XX
DE Fucosyltransferase III coding sequence.
XX
KW P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;
KW calcium-dependent carbohydrate binding protein; platelet; endothelium;
KW thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;
KW inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
KW disseminated intravascular coagulation; rheumatoid arthritis; antibody;
KW adult respiratory distress syndrome; tumour metastasis;
KW fucosyltransferase III; acetylglucosaminyltransferase;
KW autoimmune disorder; inflammatory disorder; ss.
XX
XX Homo sapiens.
XX
OS
XX
PN M09706176-A2.
XX
PD 20-FEB-1997.
XX
PF 02-AUG-1996; 96WO-US12820.
XX
PR 17-MAY-1996; 96US-0649802.
PR 03-AUG-1995; 95US-0510920.
PR 15-MAY-1996; 96US-0017794.
XX
PA (OKLA ) UNIV OKLAHOMA STATE.
XX
PI Cummings RD, Mcveer RP, Moore KL;
XX
DR WPI; 1997-154206/14.
XX
PT New O-glycan(s) derived from P-selectin glyco:protein ligand
PT (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
PT for treating inflammatory or autoimmune disorders or tumours
XX
PS Disclosure; Page 80-81; 96pp; English.
XX

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 904.255 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-4_COPY_1289_2368

Perfect score: 1080

Sequence: 1 atgacatcaacatcccaagg.....agaaatggttttgaattaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

27472414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthm:*
3: em_estln:*
4: em_estlm:*
5: em_estlv:*
6: em_estro:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estc:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	573.8	53.1	697	9	BB397507 BB397507
2	492.8	45.6	607	10	BB102760 60288338
3	422	39.1	716	9	AV725071 AV725071
4	324.8	30.1	649	9	AV338075 AV338075
5	307.4	28.5	579	10	BM021081 BM021081
6	296.8	27.5	1101	12	CNS0595ST
7	284.6	26.4	648	9	BB595164 BB595164
8	272.4	25.2	625	9	BB649967 BB649967
9	261.6	24.2	767	9	BB640911 BB640911
10	243	22.5	568	10	BM021317 BM021317
11	219.6	20.3	628	9	BB651409 BB651409
12	184.2	17.1	570	10	BM082484 BM082484
13	177	16.4	1100	12	CNS05792
14	170.2	15.8	1016	12	CNS0429P
15	165	15.3	844	12	CNS0108V
16	157.4	14.6	887	12	CNS0200S
17	155.4	14.4	651	10	BT036855 BT036855

C 18	151.6	14.0	666	10	BT050300
C 19	151.2	14.0	711	12	CNS034XF
C 20	151	14.0	501	12	B41073
C 21	146.6	13.6	640	10	BT073227
C 22	140.6	13.0	533	10	BT604066
C 23	136.6	12.6	670	9	BB626126
C 24	136.2	12.6	199	9	BB398103
C 25	123.4	11.4	674	10	BE605503
C 26	120.6	11.2	971	12	CNS0471V
C 27	120.4	11.1	615	10	B1326811
C 28	117.8	10.9	1023	12	CNS02ANG
C 29	107.8	9.5	487	10	BE63396
C 30	100.2	9.3	702	10	BT043578
C 31	93.8	8.7	863	12	CNS02ROR
C 32	90.4	8.4	720	10	BF180380
C 33	90	8.3	551	10	BT035701
C 34	87.4	8.1	601	10	BT036679
C 35	87.4	8.1	605	10	BT028804
C 36	87	8.1	1056	12	CNS01VL7
C 37	86.8	8.0	304	9	A0081478
C 38	84.2	7.8	638	10	BT035697
C 39	83.6	7.7	1057	12	CNS03ADC
C 40	83	7.7	566	10	BT090565
C 41	79.4	7.4	776	12	CNS03R10
C 42	78	7.2	838	10	BT432356
C 43	76.4	7.1	655	10	BT6538740
C 44	75	6.9	878	12	CNS02VID
C 45	71.6	6.6	590	10	BM082755

ALIGNMENTS

RESULT 1
LOCUS BB397507
DEFINITION BB397507 RIKEN full-length enriched, ES cells Mus musculus CDNA clone C330004C19 3' similar to AB014426 Mus musculus Fuf9 mRNA for alpha1,3-fucosyltransferase IX, mRNA sequence.

ACCESSION BB397507.2 GI:16411808
VERSION BB397507
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 697)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,U., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 15, 2000 this sequence version replaced gi:9216903.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

TITLE

JOURNAL

COMMENT

wasgi,K., Fujisake,S., Inoue,K., Togawa,Y., Tawa,M., Ohara,E., Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

S., Kawal, T., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y., RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kondo, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y., Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, D., Shibata, K. and Hayashizaki, Y., Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. *Genome* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
SOURCE

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C330004C19"
/clone_idb="RIKEN full-length enriched, ES cells"
/cell_type="ES cells"
/lab_host="SOLR"
/note="Site 1: XhoI site; 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAAGAGACAGACATCCAGACGCTCTTTTCTTTTTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAAGAGACATCTCGGATTAATTAATTAATTCCTCCCTCCCTCC 3'].

```

Query Match	53.1%	Score 573.8	DB 9	Length 697
Best Local Similarity	91.4%	Pred. No. 7e-14%		
Matches 619	Conservative 0	Mismatches 57	Indels 1	Gaps 1
QY	404	AATTGGAAATACACAACTCACACACACCCCAAAAAGCTGGCAATWAGACACCTGGTTTAACCGA	463	
Db	4	AATTGAGACACACACCTACACACCCCAAAAAGAGTGGCA-TGAACACTGTTCACCACTGA	62	
QY	464	CTGTGACTTACCGCGCTGATTCAGATATCCAAAGTCCCTTATGCTCTTTGAGCGGTAAAGCA	523	
Db	63	CTGTAACTTATTCGCGCTGATTCAGATATCCAAAGTCCCTTATGCTCTTTGAGCGGTAAAGCA	122	
QY	524	CAAAATCCCTGGTGTGGAAGTCCCAAGAAAGAAATGGTGTGCTGGGTGTGGAGTA	583	
Db	123	CAAAATCCCTGTGTGTGGAAGTCCCAAGAAAGAAATGGTGTGCTGGGTGTGGAGTA	182	
QY	584	ACTGGAAACCCCTGACACATGCCAGAGTCAAGTATTACATATGAGCTTAAAGCAAAAGACTTGAA	643	
Db	183	ACTGGAAACCCCTGACACATGCCAGAGTCAAGTATTACATATGAGCTTAAAGCAAAAGACTTGAA	242	
QY	644	TCCATATCTTACGAGCGCAAGCAATTTGGAGATATGTCAATGATAAATAATTTGATTCCTCA	703	
Db	243	TCCACACCCATATGGCCAAAGCAATTCGGAGAAATACGTGAACGATAAATAATTTGATTCCTCA	302	
QY	704	TATTCGCTGTGAATTTTATATCTTCTCTTTGAAAATTCATCCAAAGATATACATACGG	763	

Db	303	TATCTACTGTGAATTTTATCTTTCATTGTGAAATCAATTCACAAAGATTTACATCCAG	362
Qy	764	AAAACTCTAATAAGCTTTTCTGGCGCTGTACTGTGTGTGGAGACATCTAGGG	823
Db	363	AAAACCTCTCAATGCACTTTTGGCTGGGTGAGTACCTGTGTCTCGGGTCCATCTAGGG	422
Qy	824	AAAATATAGAAATATATATCCAGACAGATTCATTCATTATGTAATGGAATTTAACTCTC	883
Db	423	AAAACATAGAAATATATATTCACGCTGATCATTCATTCATCTCAATGGAGATTTTAACTCTC	482
Qy	884	CCAGTAGACTGCAAGATATCTGAAGGAAGTCGACAAAACAAATTAAGTTATTAACCTTAGTT	943
Db	483	CCAGTAGATTTAGCAAAATATCTGAAGGAAGTTGACAAAACAAATTAAGTTTAACTTAGTT	542
Qy	944	ACCTTTAAGTAGGAAGAGATTTCACTGTAATCTCTCCACGATTTTGGGAATACATCAT	1003
Db	543	ACCTTTAAGTAGGAAGAGATTTTACTGTAATCTCTCCACGATTTTGGGAATACATCAT	602
Qy	1004	GTTTGCTGGGATCATGTGAAGAGCATTCAGAAATTAAGTGTGTGGTAATTTAAGA	1063
Db	603	GGCTGGCATGGACCATCTATAAAGGCATCAAGAAATTAAGTGTGTGGTAATTTAAGA	662
Qy	1064	AATGGTTTGGAAATTA 1080	
Db	663	AATGGTTTGGAAATTA 679	

DESIGN

RESULT	2
LOCUS	BI102760
DEFINITION	B102760 607 bp mRNA linear EST 26-JUN-2001
ACCESSION	60288833BFF1 NCI_CGAP_Kid14 Mus musculus CDNA clone IMAGE:5043609
VERSION	5'', mRNA sequence.
KEYWORDS	BI102760 BI102760.1 GI:14553653
SOURCE	EST.
ORGANISM	house mouse. Mus musculus.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 607) NIH-GSC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC). Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D. Email: grabsb-r@mail.nih.gov
JOURNAL	Tissue procurement: jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LNL) Clone distribution: MGC Genome Genomics, Inc. found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
COMMENT	Plate: FLAM1119 row: 1 column: 10 High quality sequence stop: 606.

FEATURES

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone IMAGE:5045609"
/clone_id="NCI_CGAP_kid14"
/lab_host="DH10B (TI phage-resistant)"
/site="Organ: Kidney; Vector: PCMV-SpORe6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      179 a.      149 c      121 g      157 t
ORIGIN

```

Query Match	45.6%;	Score 492.8;	DB 10;	Length 607;
Best Local Similarity	90.1%;	Pred. No. 6.5e-123;		
Matches 549; Conservative	0;	Mismatches 58;	Indels 2;	Gaps 2;

QY 223 CAGACCTTTGACCTTACATCTGCGCAAGCATGTTCACATCCAGAGGATCCATCTCACA 292
|||||
Db 1 CAGACCTTTGAGCTTACATCTGCGCAAGCATGTTCACATCCAGAGGATCCATCTCACA 60
QY 283 ACGACCGTCTACTCTGACAAATCCCAATGCGATCTCTGATCATCAGCAGATCAGT 342
|||||
Db 61 ACGACCGTCTCA-TGTACAAATCCCAATGCGATCTCTGATCATCAGCAGATCAGT 119
QY 343 TGGGATCTGACAAATTTACTCTGAGCAGATGAGCCACCTTCAGAAATGATTTGATG 402
|||||
Db 120 TGGGATCTGACAAATTTACTCTGAGCAGATGAGCCACCTTCAGAAATGATTTGATG 178
QY 403 AATTTGGAATCCACCACTCAGCTCCCAAAAGATGGCATTTGAGACCTTTTAACTG 462
|||||
Db 179 AATTTGAGTCAACCCACTCAGACCCCAAAAGATGGCATTTGAGACCTTTTCAACTG 238
QY 463 ACTGTGACTTACCGCGGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 522
|||||
Db 239 ACTGTGACTTACCGCGGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 298
QY 523 ACAAAATCCCTGCTGTTGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 582
|||||
Db 299 ACAAAATCCCTGCTGTTGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 358
QY 583 AACTGTGACCTTGCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 642
|||||
Db 359 AACTGTGACCTTGCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 418
QY 643 ATCCATACCTTACCGGCAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 702
|||||
Db 419 ATCCATACCTTACCGGCAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 478
QY 703 ATATCTGCTGTAATTTATCTCTTCTTGAATTCATTCAGATTCAGATTCAGATTCAGAT 762
|||||
Db 479 ATATCTGCTGTAATTTATCTCTTCTTGAATTCATTCAGATTCAGATTCAGATTCAGAT 538
QY 763 GAAAACCTTACCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822
|||||
Db 539 GAAAACCTTACCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
QY 823 GAAAACCTTAT 831
|||||
Db 599 GAAAACCTTAT 607
|||||

RESULT 3
AV725071 716 bp mRNA linear EST 16-OCT-2000
LOCUS AV725071 HTC Homo sapiens cDNA clone HTCLB02 5', mRNA sequence.
DEFINITION AV725071
ACCESSION AV725071
VERSION AV725071.1 GI:10830091
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 716)
Gu Y., Peng Y., Song H., Huang Q., Yang Y., Gao G., Xiao H., Xu X.,
Li N., Qian B., Liu F., Qu J., Gao X., Cheng Z., Xu Z., Zeng L., Xu
S., Gu W., Yu Y., Jia J., Fu G., Ren S., Zhong M., Lu G., Hu R.,
Chen J., Chen Z. and Han Z.
Homo sapiens cDNA HTC clones
Unpublished (2000)
Contact: Zeguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzeghgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

TITLE
JOURNAL

source 1. 716
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTCLB02"
/clone_lib="HTC"
/tissue_type="hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 219 a 113 c 133 g 249 t 2 others

ORIGIN

Query Match 39.1%; Score 422; DB 9; Length 716;
Best Local Similarity 100.0%; Pred. No. 1.1e-103;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 659 AACCATTTGGAGATATGTCATGATTAATAATTTGATCTCTACCATATCTGCTGTAAT 718
|||||
Db 1 AACCATTTGGAGATATGTCATGATTAATAATTTGATCTCTACCATATCTGCTGTAAT 60
QY 719 TTATCTTCTCTTGAATTTCAATCCAGAGATTAATCAGAGAAAGCTTATCAATG 778
|||||
Db 61 TTATCTTCTCTTGAATTTCAATCCAGAGATTAATCAGAGAAAGCTTATCAATG 120
QY 779 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838
|||||
Db 121 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 839 ATATTCAGACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 898
|||||
Db 181 ATATTCAGACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 240
QY 899 AGATTCGAGGAGAGTCGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 958
|||||
Db 241 AGATTCGAGGAGAGTCGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
QY 959 AGGATTCACCTGTAATCTTCCAGATTTGGGAATCAGATTCAGATTCAGATTCAGATTC 1018
|||||
Db 301 AGGATTCACCTGTAATCTTCCAGATTTGGGAATCAGATTCAGATTCAGATTCAGATTC 360
QY 1019 ATGTGAAAAGGCAATCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078
|||||
Db 361 ATGTGAAAAGGCAATCAAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 1079 AA 1080
|||||
Db 421 AA 422
|||||

RESULT 4
AV338075 649 bp mRNA linear EST 24-OCT-2001
LOCUS AV338075
DEFINITION AV338075 RIKEN full-length enriched, adult male olfactory bulb Mus
musculus cDNA clone 6430407H24 3' similar to AB015426 Mus musculus
mRNA for alpha1,3-fucosyltransferase IX, mRNA sequence.
ACCESSION AV338075
VERSION AV338075.2 GI:16396626
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 649)
Arakawa T., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A.,
Hiramoto K., Hori I., Ishii Y., Ito M., Kawai J., Kono H., Kouda
M., Koye S., Matsuyama T., Miyazaki A., Nomura K., Ohno M.,
Okazaki Y., Okido T., Saito R., Sakai C., Sakai K., Sano H., Sasaki
D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H.,
Tagami M., Tagawa A., Takahashi F., Takeda Y., Tanaka T., Toyama T.,
Muramatsu M. and Hayashizaki Y.
RIKEN Mouse ESTs (Arakawa T., et al. 2001)
Unpublished (2001)

TITLE
JOURNAL

Db	408	CCAGCTCACAAGTCTCTACGTCTGATTTTGGAAATCCATCCACAAAGACTACTTATTACGTGAAA	349
QY	767	AGCATATCAATGGTTTTTTTTCGGCGGCTGTCACCTGTTGTTTCGGAGCAACATGAGGAAA	826
Db	348	AACGTATCAACACCCACTTGGCTGTGGGAGACAGTGTCCAGTGGTTCTGTGGCCACCTCGGGAGA	289
QY	827	ACATATGGAATTTATTTCCAGAGAGATCATTTATTCATTCATGAGGAAATATTAATCTCCCA	886
Db	288	ACTACGAGAACATCATCATCCACAGGATGCTTTATTCACAGTGTGAAGACTTTGGCTCACCCA	229
QY	887	GTGACGTAGCAAGTATTCGAAAGAGACGACGCAAAAACAATACTTATACCTAGTACT	946
Db	228	AGGAGCTGTGCTACTTTTACTGTCTCTCTGACCAAAAATGAGGAATGTACTTAAATTT	169
QY	947	TTTAACTGAGAGAGATTTCTACTGTAAATCTTCCACCAATTTTGGGAAATCAACGATGTT	1006
Db	168	TTTCANGTGGCGAGACACTTTAAAGCAAGGCGCATGTGTGCCAGGCG -CACCTG	110
QY	1007	TGCGTTGCCATCATCTGTGAAAAGGCATCAAGAATATA -AGTCTGTGTGTAATTTAGAGAAA	1065
Db	109	AGGCTTTGTGTTTACGTACGAGAGGCGNTNMCATATACAGTATTCAGTACGACCTTAAACAAG	50
QY	1066	TGCTTTTGGAAATTA	1080
Db	49	TGCTACTGGGGTTGA	35

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome* 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Location/Qualifiers

1. 648

```

/organism="Mus musculus"  

/strain="C57BL/6J"  

/db_xref="taxon:10090"  

/clone="C030004N01"  

/clone_1lb="RIKEN full-length enriched, adult male corpus  

striatum"  

/sex="male"  

/tissue_type="corpus striatum"  

/dev_stage="adult"  

/lab_host="DH10B"  

/note="Site_1: SalI; Site_2: BamHI; cDNA library was  

prepared and sequenced in Mouse Genome Encyclopedia  

project of genome Exploration Research Group in Riken  

Genomic Sciences Center and Genome Science Laboratory in  

RIKEN Division of Experimental Animal Research in Riken  

contributed to prepare mouse tissues. 1st strand cDNA was  

primed with a primer [5'  

GAGGAGAGAGATCCCAAGACGCTTTTTTTTTTTTNN 3'] , cDNA was  

prepared by using trehalose thermo-activated reverse  

transcriptase and subsequently enriched for full-length b  

cap-trapper.cDNA went through one round of normalization  

to Rot -10.0 and subtraction to Rot = 185.0. Second  

strand cDNA was prepared with the primer adapter of  

sequence [5' GAGAGAGATTCGTGGATTAAATTAATATCCCCCCCCCCC  

3']. cDNA was cloned into the XhoI and BamHI sites.  

Vector: a modified pluscript KS(+) after bulk excision  

from lambda FLC I. Cloning sites, 5' end: SalI; 3' end:  

BamHI"
```

RIKEN Mouse ESTs (Araiawa, T., et al., 2001)
 Unpublished (2001)
 On Nov 30, 2000 this sequence version replaced g1:11491766.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 url: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 , M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyura
 S., Kawai, Y., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 T. and Hayashizaki, Y.

1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resseqsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, K., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagik, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wathiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, Y., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shigaawa, A., Saito, T., Kiyosawa, H., Yamakura, I., Aizawa, K. and
K., Fukuda, S., Hara, A., Itoh, A., Itoh, Y., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

```

FEATURES
source
location/Qualifiers
1..767
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A730093G08"
/clone_id="RIKEN full-length enriched, 7 days neonate cerebellum"
/tissue_type="cerebellum"
/dev_stage="7 days neonate"
/lab_host="DH10B"
/note="Site.1: SalI; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGAGAGAAGATCCACAGACCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adaptor of sequence [5', GAGGAGAGAGATTTCGACTGTAAATTAATATCCCCCCCCCCC 3'] cDNA was cleaved with XhoI and BamHI. Vector: a modified p Bluescript KS(+)) after bulk excision from Lambda
FLC I.a
BASE COUNT      175 a      239 c      176 g      172 t      5 others
ORIGIN
```

	Query Match	261.6;	Score	261.6;	DB	9;	Length	767;	
	Best Local Similarity	88.0%;	Pred.	No. 4.1e-60;					
	Matches	285;	Conservative	0;	Mismatches	39;	Indels	0;	Gaps
OY	1	ATGACATCAACATCCAAAGGAATTCTTGCGCCCATTTTAAATGTGTCGATTAATCCGGGC	60						
Dd	347	ATGACATCAACATCCAAAGCATTCTTGCCCATTTTAAATGTGTCGATTAATCCGGGC	406						
OY	61	TGTTTCATGGCATGTGTTCTTCATTCATCAACAACCTACCACAGGTGATCTTCAGTCCA	120						
Dd	407	TGTTTCATGGCATGTGTTCTTCATTCATTCATCAACAACCTACCACAGGTGATCTTCAGTCCA	466						
OY	121	ATGGAATCAGCCAGCTCTGTGCTGGAATAATGAAAAACTTTTTCCACCAAAACTGATTAAT	180						

Accession	Sequence	Position
Db	ATGAGAGCTGGCAAGCTTCTGGGGCTCAAAATGAAATATTTCTCTCCACCAAAACGATATAT	526
Qy	TTTATGAAACTACTATTCTGTGTGGGTGTGGCAATTGGGCAAGACCTTTGACCTTACA	240
Db	TTTAAACGAAGCTACCATCTCTGTGTGGGTATGGCCATTTTGGCAAGACCTTTGACCTTTACA	586
Qy	TCTGTCGCCAGCAAGTTTACATCCAGAGATGCCATCTCATAACAGGAAACCCCTTCACCTGAC	300
Db	TCTGGCCAGCAAGATTTTAAATATTCAGGGGGGCCATCTTACACAAAACGGTCACTGAC	646
Qy	AACAAATGCCATGCATTTCTGATC	324
Db	AACAAATGCCAGGGGGGCGCTGATC	670

RESULT 10	
BM021317	
LOCUS	
DEFINITION	BM021317 568 bp mRNA linear EST 30-Oct-2001
	ie75f03.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
	cDNA 5', similar to TR:Q9Y231 Q9Y231 ALPHA-3-FUCOSYLTRANSFERASE.1.
	mRNA sequence.
ACCESSION	BM021317
VERSION	BM021317.1
KEYWORDS	EST.
SOURCE	Human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 568)	Melton, D., Brown, J., Kenly, G., Permutt, A., Lee, C., Kastner, K., Hillier, L., Scarsie, M., Bresnelli, J., Grawohn, G., Clifton, S., Miller, L., Motta, M., Pape, D., Willet, T., Martin, J., Bisthan, A., Smith, A., Theising, B., Ritter, E., Konno, T., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarelisvili, R., Williams, I., Jackson, Y. and Bowers, Y.	Endocrine Pancreas Consortium	Unpublished (2000)	Other-ESrs: le75f03:xl

Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownjef@wustl.edu)
High quality sequence stop: 436.

```

FEATURES
SOURCE
Location/Qualifiers
1..568
/Organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue-type="Islets of Langerhans"
/digest="Age"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPOR1; Site:1: Not 1;
Site:2: Sal 1; Starting library constructed using
SuperScript plasmid library kit (Life Technologies) . cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:1791-806; 0.5 micrograms single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot
20. Single-stranded (unhybridized) plasmids were isolated

```

by hydroxyapatite chromatography and used to make this library."

BASE COUNT	127 a	170 c	123 g	148 t
ORIGIN				

Query Match	22.5%	Score 243	DB 10	Length 568
Best Local Similarity	100.0%	Pred. No. 4	1e-55	
Matches 243	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	ATGCATCAACATCCAAAGGAATTTTGGCCATTTTAAATGTGCTATTAATTCGGGC	60
Db	326	ATGCATCAACATCCAAAGGAATTTTGGCCATTTTAAATGTGCTATTAATTCGGGC	385
QY	61	TGTTTCATGCGATGCTTCTCATTTATTCATCAAACTACCACAGCTGATCTTCAGTCA	120
Db	386	TGTTTCATGCGATGCTTCTCATTTATTCATCAAACTACCACAGCTGATCTTCAGTCA	445
QY	121	ATGGAATAGCGCAGCTGCTGCTGAAATAATGAAAACTCTTTCCACCAAACTATTTAT	180
Db	446	ATGGAATAGCGCAGCTGCTGCTGAAATAATGAAAACTCTTTCCACCAAACTGATTTAT	505
QY	181	TTTAAATGAACACTACTATTCTGCTGCTGCGCTGCGCCATTTGGGCAGACCTTTGACCTTACA	240
Db	506	TTTAAATGAACACTACTATTCTGCTGCTGCGCTGCGCCATTTGGGCAGACCTTTGACCTTACA	565
QY	241	TTC 243	
Db	566	TTC 568	

LOCUS	DEFINITION	EST
BB651409	628 bp mRNA linear	26-OCT-2001
BB651409	RIKEN full-length enriched, 0 day neonate	cerebellum Mus
musculus cDNA clone C230079N02 5', mRNA sequence.		

REFERENCE	AUTHORS	TITLE
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 628)	Arakawa, T., Carlinici, P., Fukuda, S., Furuno, M., Hanaagaki, T., Hara, A., Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, R., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shioda, K., Shingawa, A., Shirai, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

mRNA sequence
D90874.4

ACCESSION	BM082484
VERSION	BM082484.1
GI	GI:16929414

BM082484.1 GI:16929414

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE	JOURNAL	COMMENT
Washu zebrafish EST Project 1998	unpublished (1998)	Other_ESTs: fu25e08.y1

Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbratf@watson.wustl.edu
cDNA Library constructed by library constructed by Invitrogen and
donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution by: The I.M.A.G.E. Consortium/LLNL, send email to:
info@imgl.llnl.gov
High quality sequence stop: 419.

FEATURES	Location/Qualifiers
source	1. .570

```

/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5306223"
/clone_1lb="Campbell zebrafish ovary"
/sex="female"
/dev_stage="4-5 months, 1 year and 2 years"
/lb_host="DH108 (phage-resistant)"
/note="Organ: ovary(pooled); Vector: pCMV-Sport6; Site:1;
Nolt: Site:2; Salt: Whole ovaries collected from zebrafish
aged 4-5 months, 1 year and 2 years. Oligo-dt primed,
directionally cloned. Average insert size 2 kb. Library
constructed by Invitrogen and donated by R. Campbell
(Marine Biology Laboratory, Woods Hole, MA)."
```

[illegible]

QY 673 TATGTCATATATATAAATTTGATTCCTACACATATCGCTTAAATTTTATCTTCCTT 732
 Db 270 TATCTACATATATAAAGACTACTATCTTACCCCGCTACTCCAAATTTTACCTCTCTTC 211
 QY 733 GAAATATCATATCCACAGAGATTACATATCACAGAAAAGCTATACATATGCTTTTCTGGCGGC 792
 Db 210 GAGAGCCCTCCACAGAGACTACATACACTGAGAAAGATCATATGCTCTCTTGGTCAGAGA 151
 QY 793 TCTGATACCTTTTCTTGGGACCATCTATGGGAAAACTTGTGAAATTTATATTCAGACAGAA 852
 Db 150 ACGTTCTCTGTGTTTGGCTCTCCCAAGAAAAACATGAAAACTTCATCTCTGGAGAC 91
 QY 853 TCATATCTCATSTGGAAATATATACATCTCCGACAGTAGCTGAGAAAGTATCTGGAAGGAA 912
 Db 90 GCCCTTATTTCTATSTGAAATGACTTCTCTAAGTGTGCAAGTCTCTACACACATATTTGTATAG 31
 QY 913 CTCGACAAAACAATATAGTTATACCTTAG 941
 Db 30 CTAATATATAAATGACATTCATATACAG 2

RESULTS	LOCUS	DEFINITION
13	CNS05792/c	1100 bp DNA linear GSS 26-JUL-2000
13	CNS05792	Tetradon nigroviridis genome survey sequence T7 end of clone

VERSION	AL324368.1	GI:9557250
KEYWORDS	GSS; genome survey sequence	
SOURCE	Tetraodon nigroviridis.	
ORGANISM	Tetraodon nigroviridis	

REFERENCE

AUTHORS

TITLE	JOURNAL	REFERENCE	AUTHORS
Estimate of human gene number provided by genome-wide analysis using tetraodon nigroviridis DNA sequence	Nat. Genet.	25 (2), 235-238 (2000)	20296633
2 (bases 1 to 1100)			
Crollius,H.R., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Bernot,A., Fizames,C., Mincker,P., Brotlier,P., Quetier,F., Sautin,W. and Weissenbach,J.			

TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish tetraodon nigroviridis
JOURNAL MEDLINE	Genome Res. 10 (7), 939-949 (2000)
REFERENCE	20359837
AUTHORS	3 (bases 1 to 1100)
TITLE	genomecope.
JOURNAL	Direct Submission
COMMENT	Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .
FEATURES	Location/Qualifiers

BASE COUNT	314 a	228 c	252 g	304 t	2 others
ORIGIN	/note="Genoscope sequence ID : C0AA007CD03C1-end : T77"				

Query Match	16.4%	Score 177;	DB 12;	Length 1100;
Best Local Similarity	55.4%	Pred. No. 4.2e-37;		
Matches 428; Conservative	0;	Mismatches 337;	Indels 13;	Gaps 4;

OY	785	TGGCTGGCTGTCGTACCTTGTGTGGAGCACCATGAGAAACATGAGAATTAT-T	843
Db	116	TGGCTGGCTGTCGTACCTTGTGTGGAGCACCCTCCAAAAAATAATTTAT	57
OY	844	CCAGCATTCATTCATTCATCGTAAT	874
Db	56	CCTTGCAATTCGTTATTCAGCTTAACACT	26

RESULT 15

CNSOLV8Y			
LOCUS			
DEFINITION	CNSOLV8Y	844 bp	DNA linear GSS 12-MAY-2000
	Tetradon nigroviridis genome survey sequence PUC-Orig end of clone		
	19K09 of library G from Tetradon nigroviridis, genomic survey		
	sequence.		
ACCESSION	AL168808.1	GI:7806865	
VERSION	GSS: genome survey sequence.		
KEYWORDS	Tetradon nigroviridis.		
SOURCE	Tetradon nigroviridis		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthomorpha; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthopterygii; Acanthopterygii; Percomorphi; Tetraodontiformes; Tetraodontidae; Tetradon.		
REFERENCE	1 (bases 1 to 844)		
AUTHORS	Roeest-Croilius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.		
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 844)		
AUTHORS	Roeest-Croilius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J.		
	Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence		
TITLE	Unpublished		
JOURNAL	Tetradon nigroviridis		
REFERENCE	3 (bases 1 to 844)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases		
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .		
FEATURES	Location/Qualifiers		
SOURCE	1..844		
	/organism="Tetradon nigroviridis"		
	/db_xref="taxon:99883"		
	/clone="198X09"		
	/clone_1id="G"		
	/note="Genoscope sequence ID : C0AG198AF05SP1-end ; PUC-Orig"		
BASE COUNT	199 a 190 c 205 g 245 t	5 others	
ORIGIN			
	Query Match	15.3%; Score 165; DB 12; Length 844;	
	Best Local Similarity	54.1%; Pred. No. 7e-34;	
	Matches 399; Conservative	3; Mismatches 324; Indels 12; Gaps 3;	
OY	20	GAATTCCTCGGCCATTTTATTCGTCATTCACCTTCGCGGCTGTTTCATGCATGCTCWC	79
Db	119	RAACATCGCGCCCTTTTCATTCGCTGTTTTGTTTTGGGGCCCTGTGCACTTTAATCT	178
OY	80	TCATTTCATCAACACCTTCCAACAAGCTGACATTCAGTCCAAATGGAATCAGCACCTCTG	139
Db	179	TTTAGCTTTGGAACCAATCCACTGCTGCGCTTAAAGTCTGCTGGAACA-----ACAG	232
OY	140	TGCTGAAAATGAAAATCTTTTTCACCAAAACTGATTTATTTAATGAACACTACTATTC	199
Db	233	CAGCAGCCGTATAGAGAAATCATTTCCGACAGACGGATAA--AAACCTGACCAACGGTGC	289

Oy	200	TGGTGTGGGTGGCCATTGTTGGGACAGCTTTGACCTTAACTCCGTGCAAGCAATGTGCA	259
Db	280	TGCTGTGGTTCGGCAATTTGGGATTTACCTTTGACCCGAAAGTGTGAGCGCTTACTATG	349
Oy	260	ACATCCAAAGATATGCATCTCCACAGGAGCGGTTCACTGTAACAACAAATCCCATGCAATTC	319
Db	350	GTAATAGAGCGGTGTTTCAATTCACAGCGGACAGGAACATTTCAAGGAAGAAATGCGNTCG	409
Oy	320	TGATCATCACCAGGACATCAGTGGGATCTGCAAATTTAACTCCAGCAGATAGGCGAC	379
Db	410	TTATTCATCATCGAAGAAATTTCTCCGACACTGTCCATCTGCGGACACAGACGCGCGT	469
Oy	380	CCCTTCGAAATATGATTTGGATTTGGAAATTTGGAAATCCCAACTCCACACGCCCAAAACAGT	439
Db	470	TGTACCAAAATGTGGGTCTGGTTCAACATGAGAAATCCTTGTACTCCGACAGGTTCTTG	529
Oy	440	GCAATGAGCACTGTTTAACTGACTGCTGACTTACGCGCGTATGTAGATATACCAAGTC	499
Db	530	GAAATGAGCGGTTGTTCACCTGACTGTGAATTTGCTGTGATGCTGATATTTTCCATGC	589
Oy	500	CTTATGCGTTCTTGACGTTAGCAACAATCCCTTCTGTTGAAGTCCCAAGCAAGAGA	559
Db	590	CTCAATGTACTGTTCTCTGAAAAAAGTAGAAGACATTGTGCCCCCAGCAAAAGATA	649
Oy	560	AATTTGCTGCTGGCTTGATGAGTAACTGCAACCTGTAGCATGCCAGATCAAGTATTTACA	619
Db	650	AGTTGGTCTGCTGGATATCTACGCAACATGGAAGAGCACTTTATATAGGGGTAAATATCTCA	709
Oy	620	ATGAGCTTAGAAAAGCAATTGAATTCATATCTTACGGGCAAGCAATTTGGAGAAATATGTCA	679
Db	710	AGGAGCTGTCCCAAGCACATTAAGATACCACTTATAGGACGTGGCTTTGGGGGCTA--CA	766
Oy	680	ATGATAAAATTTGATTCATCCATATCTGCTGTAAATTTATCTTCCCTTTGAAAATTT	739
Db	767	ACTGGGGAAGCGCTCGTCACTCATATGGCCAGCTGTAAATGTTACCTGTCTTTTGAAGAACT	826
Oy	740	CAATCCACAAGATTTACA	757
Db	827	CCATCTACATTTGACTACA	844

Search completed: October 6, 2002, 07:16:06
Job time : 912.255 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 22.0361 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-4_COPY_1289_2368

Perfect score: 1080
Sequence: 1 atgacatcaacatccaaagg.....agaatggtttggaattaa 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	11.8	1814	2	US-08-483-151-1 Sequence 1, Appli
2	127	11.8	1814	5	PCT-US96-06427-1 Sequence 1, Appli
3	113.4	10.5	1654	1	US-07-914-281-13 Sequence 13, Appli
4	113.4	10.5	1654	1	US-08-393-246-13 Sequence 13, Appli
5	113.4	10.5	1654	1	US-08-525-058A-13 Sequence 13, Appli
6	113.4	10.5	1654	2	US-08-696-731-13 Sequence 13, Appli
7	113.4	10.5	1654	4	US-09-042-531-13 Sequence 13, Appli
8	112.2	10.4	1086	1	US-07-914-281-12 Sequence 12, Appli
9	112.2	10.4	1086	1	US-08-393-246-12 Sequence 12, Appli
10	112.2	10.4	1086	1	US-08-525-058A-12 Sequence 12, Appli
11	112.2	10.4	1086	2	US-08-696-731-12 Sequence 12, Appli
12	112.2	10.4	1086	2	US-09-042-531-12 Sequence 12, Appli
13	112.2	10.4	2043	3	US-09-063-237-2 Sequence 2, Appli
14	112.2	10.4	2043	1	US-07-914-281-1 Sequence 1, Appli
15	112.2	10.4	2043	1	US-08-393-246-1 Sequence 1, Appli
16	112.2	10.4	2043	1	US-08-273-411-4 Sequence 4, Appli
17	112.2	10.4	2043	1	US-08-525-058A-1 Sequence 1, Appli
18	112.2	10.4	2043	2	US-08-696-731-1 Sequence 1, Appli
19	112.2	10.4	2043	4	US-09-042-531-1 Sequence 1, Appli
20	112.2	10.4	2043	5	PCT-US91-00899-1 Sequence 1, Appli
21	111.8	10.4	1316	1	US-07-914-281-10 Sequence 10, Appli
22	111.8	10.4	1316	1	US-08-393-246-10 Sequence 10, Appli
23	111.8	10.4	1316	2	US-08-525-058A-10 Sequence 10, Appli
24	111.8	10.4	1316	2	US-08-696-731-10 Sequence 10, Appli
25	111.8	10.4	1316	4	US-09-042-531-10 Sequence 10, Appli
26	100.8	9.3	1068	5	PCT-US91-00899-13 Sequence 13, Appli
27	100.8	9.3	1215	5	PCT-US91-00899-5 Sequence 5, Appli

28	100.8	9.3	1488	1	US-07-914-281-9 Sequence 9, Appli
29	100.8	9.3	1488	1	US-08-393-246-9 Sequence 9, Appli
30	100.8	9.3	1488	1	US-08-525-058A-9 Sequence 9, Appli
31	100.8	9.3	1488	2	US-08-696-731-9 Sequence 9, Appli
32	100.8	9.3	1488	2	US-09-042-531-9 Sequence 9, Appli
33	100.8	9.3	2175	4	US-08-482-073-9 Sequence 9, Appli
34	100.8	9.3	2861	4	US-08-482-073-10 Sequence 10, Appli
35	100.8	9.3	3647	1	US-07-914-281-7 Sequence 7, Appli
36	100.8	9.3	3647	1	US-08-393-246-7 Sequence 7, Appli
37	100.8	9.3	3647	1	US-08-525-058A-7 Sequence 7, Appli
38	100.8	9.3	3647	2	US-08-696-731-7 Sequence 7, Appli
39	100.8	9.3	3647	4	US-09-042-531-7 Sequence 7, Appli
40	100.8	9.3	3647	5	PCT-US91-00899-4 Sequence 4, Appli
41	97.6	9.0	2134	2	US-08-483-151-3 Sequence 3, Appli
42	97.6	9.0	2134	5	PCT-US96-06427-3 Sequence 3, Appli
C 43	37.2	3.4	1728	1	US-08-403-866-12 Sequence 12, Appli
C 44	37.2	3.4	2231	1	US-08-403-866-14 Sequence 14, Appli
C 45	37.2	3.4	12720	1	US-08-403-866-11 Sequence 11, Appli

ALIGNMENTS

```
RESULT 1
US-08-483-151-1
; Sequence 1, Application US/08483151
; Patent No. 5858752
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Holgersson, Jan
; TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,151
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/278001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-483-151-1

Query Match 11.8%; Score 127; DB 2; Length 1814;
Best Local Similarity 52.3%; Pred. No. 3e-28;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

QY 269 GATGCACCTCCTCAACGACCGCTTCACGTCATCAACAAATCCATGCAGTTCGATGCATC 328
DB 548 GCTGCGCTGTGAGTCTACCGACCGCTCTGACGACGTCGATGCTGTGTGCTTCACCC 607
```

Tue Oct 8 10:18:27 2002

us-09-744-748-4_copy_1289_2368.rml

Page 2

```
QY 329 ACCGAGACATCAGTTGGATCTGACAAATTTACCTACGACAGCTAGGCCCTTCCAGA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 ACCGTGACCTGCAAAACCCGCAATCTCTCCATCCCTGGACGACGAGGACGAGCAGC 667
QY 389 AATGATTTGGATGAATTTGGAAATCCCACTACACTCCCAAGAGAGGATGAGC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CTGGGTGCTGGGCTCCATGAGAAATGCGCCATTAATCCATGCTCCATGCTCCGG 727
QY 449 ACTGTTTAACCTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 GCATCTCAACTGGGTGCTGAGCTATGCGGTGATTCAGATATCTTTGACCTTACGGTC 787
QY 509 TCTTGACGGTAAAGCAAAATCCCTGCTGTTTGAAGTGGCCAAAGAAAGAAATGGGT 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 GCTTGAGAGCCTCTCTCTGGGCC---CACATCCCACTACCGGCGCAAGAGAGATGGCTG 844
QY 569 GCTGGTGTGAGTAACCTGGAACCTGAGCAGCTGCGAGAGTCAAGTATTAAGTAA 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CTTGGGTGATCAGCAATTTCCAGAGCGGAGCGAGCTGCAAGCTGACGCGAGCTGG 904
QY 629 GCAAAAGCATTTGAATCCATACCTACGAGGAGGAGGATTTGGAGATATGTAATGAAA 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 CCCCTCATCTGACAGTGTGATGTGCTGCGTCCGACGAGGAGGAGGAGGAGGAGTAT 964
QY 689 ATTGATTCCTACATATCTGCTGTAATTTATCTTCTTGAATTTCAATTCACA 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 GTCTGGTGGCCACTTTGGCCCGGTACGCTTCTACGCTGGGCTTTGAGAACTACAGCATC 1024
QY 749 AGGATTTACATCAGGAAAGCTAT---ACAAATGCTTTCTGCGTCTGCTGACTGCTG 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 GGGACTTCATCACTAGAAAGTTCTGGGCAATGCGCTGGGAGCTGTGTACCCGTTGG 1084
QY 806 TTCTGGGACATCTAGGAAACCTATGAAATTAATTCACGACATTCATTCATTCATG 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 CGGTGGACCTCTCTGGGACCTACGAGGCTTTTGTGCGCACAGATGCTTTGTACAG 1144
QY 866 TGAAGATTTAATCTCTCCAGTGAAGTAC 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1145 TGAAGATTTAATCTCTCCAGTGAAGTAC 1175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 2
PCT-US96-06427-1
: Sequence 1, Application PC/TUS9606427
: GENERAL INFORMATION:
: APPLICANT: The General Hospital Corporation
: TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/06427
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/483,151
: FILING DATE: 07-JUN-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Lech, Karen F.
: REGISTRATION NUMBER: 35,238
: REFERENCE/DOCKET NUMBER: 00786/278M01
: TELECOMMUNICATION INFORMATION:
```

```
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1814 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: PCT-US96-06427-1
```

```
Query Match 11.8% Score 127; DB 5; Length 1814;
Best Local Similarity 52.3% Freq. No. 3e-28;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
```

```
QY 269 GATGCCATCTCAACAGGAGCCCTTCACTGATACAAATCCCATGCTGATCCATC 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 GCTGCCGTGTAGTGTCTAACCCGAGCCTGTAGCCAGTGTGATGCTGTGCTTCCACC 607
QY 329 ACCGAGACATCAGTTGGGATCTGACAAATTTAAGTACAGTACGAGCCACCTCCAGA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 ACCGTGACCTGCAAAACCCGCAATCTCTCCATCCCTGGAGAGGAGGAGGAGGAGCAGC 667
QY 389 AATGATTTGGATGAATTTGGAAATCCCACTACACTCCCAAGAGAGGATGAGC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CTGGGTGCTGGGCTCCATGAGAAATGCGCCATTAATCCATGCTCCATGCTCCGG 727
QY 449 ACTGTTTAACCTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 GCATCTCAACTGGGTGCTGAGCTATGCGGTGATTCAGATATCTTTGACCTTACGGTC 787
QY 509 TCTTGACGGTAAAGCAAAATCCCTGCTGTTTGAAGTGGCCAAAGAGAAATGGGT 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 GCTTGAGAGCCTCTCTGGGCC---CACATCCCACTACCGGCGCAAGAGAGATGGCTG 844
QY 569 GCTGGTGTGAGTAACCTGGAACCTGAGCAGCTGCGAGAGTCAAGTATTAAGTAA 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CTTGGGTGATCAGCAATTTCCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGTAT 904
QY 629 GCAAAAGCATTTGAATCCATACCTACGAGGAGGAGGATTTGGAGATATGTAATGAAA 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 CCCCTCATCTGACAGTGTGATGTGCTGCGTCCGACGAGGAGGAGGAGGAGGAGTAT 964
QY 689 ATTGATTCCTACATATCTGCTGTAATTTATCTTCTTGAATTTCAATTCACA 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 GTCTGGTGGCCACTTTGGCCCGGTACGCTTCTACCTGCGCTTGAAGACATCACACATC 1024
QY 749 AGGATTTACATCAGGAAAGCTAT---ACAAATGCTTTCTGCTGCTGCTGCTGCTG 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 GGGACTTCATCACTAGAAAGTTCTGGGCAATGCGCTGGGAGGCTGTGTACCCGTTGG 1084
QY 806 TTCTGGGACATCTAGGAAACCTATGAAATTAATTCACGACATTCATTCATTCATG 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 CGGTGGACCTCTCTGGGACCTACGAGGCTTTTGTGCGCACAGATGCTTTGTACAG 1144
QY 866 TGAAGATTTAATCTCTCCAGTGAAGTAC 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1145 TGAAGATTTAATCTCTCCAGTGAAGTAC 1175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 3
US-07-914-281-13
: Sequence 13, Application US/07914281
: Patent No. 5324663
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B.
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
: TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
: TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
```

```

ADDRESS: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-914-281-13

Query Match      10.5%; Score 113.4; DB 1; Length 1654;
Best Local Similarity 51.2%; Pred. No. 3.2e-24;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

QY 196 ATTCTGCTGGGTGGCCATTGGGCGAGACCTTTGACCTTACATCTGCCAAGCATG 255
DB 318 ATCTGCTGTGAGCGCTTTTACAAACCATAGCTCTGCCGCTCAGAGAG 377
QY 256 TT---CAACATCCAGAGATCCATCTCACAGGACCGTCTACGTACAAACAAATCCAT 312
DB 378 GTGCTGAGCGCTGACCAACATCTGCGCAGCGAAGGTATCTCACAGGAGAC 437
QY 313 GAGTCTGATCCATCACCAGACATCAGTTGGATCTGA---CAATTTACCTCAGCAA 369
DB 438 GGGTATGTGTCACACCGAGAGATGATCAACCCAGTGCCTCCACGCTCC 497
QY 370 GCTAGGCCACCTTCCAGAAATGATTTGATGAAATTTGAAATACCAACTCAGACTCC 429
DB 498 CCGAGCGCGCAGGGGAGGATGATCTGTTACGATGAGTCCCAAGCCACTGCTGG 557
QY 430 CAAAAGAGGATTTAGAGACTTTTAACCTGACTCTGACTTACCGCCGTGATTAGAT 489
DB 558 CAGCTTAAGAGCATGAGAGATCTCAATCTCACCATCTCCACGCGACTCCGAC 617
QY 490 ATCCAGTGCCTTATGGCT-----TCTTACGCTAAGCAACAATCCCTCGTGT 540
DB 618 ATCTTACGCTCAGCGCTGGCTGGAGCGGTGCTCCGGCAGCTCCCAACCGCGCTC 677
QY 541 GAAGTCCCAAGAGAGAAATTTGTGTCTGGGTTTGAGTAAGTGAACCTGAGCAT 600
DB 678 AACCTCTCGCCCAAGACCGAGCTGGCTGGGAGTGTCCAACTGGGGGCCAAACTCC 737
QY 601 GCCAGAGTCAAGTATTAAGATGAGTAAAGCAAAAGATGAAATCATACCTAGCGGCAA 660
DB 738 GCCAGGCTGCGCTACTACAGACCTGCGAGGCCCATCTCAAGGTGAGCTGACGG--A 794
QY 661 GCATTTGAGAAATATGTCATGATTAATAAATTTGATCTTACCATCTCTGTAATTT 720
DB 795 CGCTCCCAAGGCCCTTGCCCGCAGGAGACCATGATGAGAGCGCTGTCCGGGTACAAGTTC 854

```

```

QY 721 TATCTTCTTTGAAATTCATCCAGAGATTTACATACGAGAAAGCTAT---ACAAAT 777
DB 855 TATCTGCGCTTGCAGAACTCTTGACACCCGACTACATACCGAGAAAGCTGGAGAGAC 914
QY 778 GCTTCTCGGCGCTCTGACTCTGTTGTTGCGGACCATCTAGGAAACATATGAGAT 837
DB 915 GCCCTGAGGCGCTGGCGCGCTGGCTGGGCGCCAGCAAGAAAGCACTACGAGAG 974
QY 838 TATATTCAGACAGATTCATTCATGTTGGAAGATTTATACCTCCAGTAGCTAGCA 897
DB 975 TTCTCTCCACCGCAGCCCTTCACTACAGTGGAGAGACTTCCAGAGCCCAAGACCTGGCC 1034
QY 898 AAGTATCTAAGAGAGTGCAGAAAACAAATAGTATACCTTAGTTACTTTACTGAGAG 957
DB 1035 CGGTACTCAGAGAGCTGAGCAAGACACAGCCCGCTACTGAGCTAGCTTGGCTGGCG 1094
QY 958 AAG 960
DB 1095 GAG 1097

RESULT 4
US-08-393-246-13
; Sequence 13, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; NUMBER OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-393-246-13

Query Match      10.5%; Score 113.4; DB 1; Length 1654;
Best Local Similarity 51.2%; Pred. No. 3.2e-24;

```


CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-13

Query Match 10.5%; Score 113.4; DB 4; Length 1654;
Best Local Similarity 51.2%; Pred. No. 3.2e-24;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

QY 196 ATTCGCTGGTGGTGGCCATTGGGCGAGACCTTGACCTTACATTCCTCCCAAGCATG 255
DB 318 ATTCGCTGGTGGTGGCCATTGGGCGAGACCTTGACCTTACATTCCTCCCAAGCATG 377
QY 256 TT---CAACATCCAGATGCGCATCTCACAAGGACCGCTTACTGTACCAAAATCCCAT 312
DB 378 GTGCTGGCAGCGCTGACTGCAACATCTCCGACCGCAAGGTGTATCCACAGGAGAC 437
QY 313 GCAGTCTGATTCATCACCAGACATCATGTTGGATCTGA--CAATTACCTCAGCAA 369
DB 438 GCGGTCACTGTCACACCGAGAGTGTATGCAACCCAGTCCCGCGCTCCACGCTCC 497
QY 370 GCTAGCCACCTTCAGAAATGATTTGATGATTTGATGATTTGATGATTTGATGATTTG 429
DB 498 CCGAGCGCGCGAGGCGAGCGATGATGATGATGATGATGATGATGATGATGATGATG 557
QY 430 CAAGAAGTGGCATGACATCTTTTAACTGACTGACTGACTGACTGACTGACTGACTGACT 489
DB 558 CAGGTGAAGCCATGAGAGGATTTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 617
QY 490 ATCCAAAGTCCCTTAAAGT-----TCTGACGCTTAAGCACAATTCCTTCGCTTT 540
DB 618 ATCTTACAGCCCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
QY 541 GAGTGGCAGCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 678 AACCTTGGCCCAAGACGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
QY 601 GCCAGAGTCAAGTATTCATATGACTAAGCAAAAGCAATTGAATTCATACCTACGGGCAA 660
DB 738 GCCAGGCTGCTTACTACACAGCTGACAGGCCCATCTCAAGGTGAGCGTATGAG--A 794

QY 661 GCATTTGGAGAAATATGTCATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 795 CGCTGCCAGAGCCCTTCCCGGAGACCATGATGATGATGATGATGATGATGATGATG 854
QY 721 TATCTTCTTTGAAATTCATATCCACAGATTTACATACGGAAGAT---ACAT 777
DB 855 TATCTGCTTTGAGAACTCTTTCACCGACTATACACGAGAAAGTGGAGAAC 914
QY 778 GCTTTTCTGCTGCTCTGTACCTGTTGTTCTGGACCATCTAGGAAACTATGAGAT 837
DB 915 GCCCTGGAGCCCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
QY 838 TATATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 897
DB 975 TTCTGCCACCGCAGCCCTTATCAGCTGAGACGCTTCCAGAGCCCAAGGACCTGGCC 1034
QY 898 AAGTATCGAAGGAAGTCGACAAAACATTAATTAATTAATTAATTAATTAATTAAT 957
DB 1035 CGTACCTGCAAGAGCTGAGCAAGAGACAGCCCGCTACTGACTGACTGCTGCTG 1094
QY 958 AAG 960
DB 1095 CAG 1097

RESULT 8
US-07-914-281-12
Sequence 12, Application US/07914281
Patent No. 5324663

GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OHION, SPIVAK, MCLELLAND, MATIER & NEUSTADT,
ADDRESS: P C
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-914-281-12

Query Match 10.4%; Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 6e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

QY 271 TGGCATCTACAAGGACCGTTCATCTGTAACAAATCCCATGCTTGTATCATCAC 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TGGCATCATCTACCGCCGCGCAAGGTGTACCCACAGGACAGACGGTCACTCGACACAC 330
QY 331 CGAGACATAGTGGGA---TCTGACAAATTTTACCTCAGACAGTGGCCACCCTTCCAG 387
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TGGGATATATATGTCACACCTTAAGTACGCTCCACCTTCCCGCGGCGGAGGGGCGAG 390
QY 388 AATGATTTGGATTTGATTTGGAATCACAACACTCACACTCCCAAAAGAGTGGCATTTAG 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CGGTGATCTGTTTCACTTGGAGCCACCCCTTAAGTCCAGAGACTGGAGGCTTGAGC 450
QY 448 CACTGTTTAACCTGACTGCTGACTTACCGCGCTGATTCAGATATCCAAAGTGCCTTAAGC 507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AGATATCTCAATCTCACACTGTCATGCTTACCGAGAGACTCCGACATCTTCAAGCCCTTACGGC 510
QY 508 TTCTTGACGCTAAGCACAATCCCTTCGT-----GTTGAAGTGGCAAGCAAGAG 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 TGGCTGGAGCCGCTGTCCGCGCCAGCCTGCCACCCACCTCAACCTCTCGGCGCAAGACC 570
QY 559 AATTTGTTGCTGGTGTGAGTATGTAAGTCAAGCCTGAGCATGCGAGAGTCAAGTATTAC 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GAGCTGTGGCTGGCGGCTGTCTCAACTGGAAGCCGAGCTCAGCCAGGCTGCGCTACTAC 630
QY 619 AATGACTTAAGCAAAAGCATTTGAATTCATACCTACGAGGCGAAGCATTTGGAGATATGTC 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAGAGCTGCAAGGCTCATCTCAAGGTGAGCTGACGAGTACGAGAGCTCCACAAAGCCCTGCGCC 690
QY 679 AATGATTAATAATTTGATTTCTTCACTACATATGCTGTTAAATTTTATCTTCTTGAAGAT 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 AAGGGAGCA--TGATGAGAGAGCTGTCCCGGTACAAAGTCTACCTGCGCTTCGAGAAC 747
QY 739 TCAATCCACAAGATATATACATCAAGCAAAAGCTAT---ACAATGCTTTTCTGCGCTGCTC 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 TCTTTCGACCCCGACTACATCACTCAAGAGAGCTGTGAGAGAGCCCTGAGGCTGAGGCC 807
QY 796 GTAAGCTGTTGTTCTGTGGACATCTAGGAAACTATGAGATTAATATTCAGCAGATATCA 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GTGCCGCTGTGCTGTGGGCCCCAGCAGCAAGCAACTACGAGAGGTCTGCGACCCGAGCGCC 867
QY 856 TTATTCATGTGGAATATATATACCTCCAGAGAGCTAGCAAAAGTATCTGAAGGAAGTC 915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 TTATTCATGTGGAATATATATACCTCCAGAGAGCTAGCAAAAGTATCTGAAGGAAGTC 927
QY 916 GACAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 GACAAAGACACGCGCGCTACCTGAGCTTCTGCTGGCGGAG 972

```

RESULT 9
US-08-393-246-12
Sequence 12, Application US/08393246
Patent No. 5595900

GENERAL INFORMATION:
APPLICANT: LOME, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLION, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-393-246-12

```

Query Match 10.4%; Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 66-24; Mismatches 323; Indels 18; Gaps 4;

```

QY 271 TGGCATCTACAAGGACCGTTCATCTGTAACAAATCCCATGCTTGTATCATCAC 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TGGCATCATCTACCGCCGCGCAAGGTGTACCCACAGGACAGACGGTCACTCGACACAC 330
QY 331 CGAGACATAGTGGGA---TCTGACAAATTTTACCTCAGACAGTGGCCACCCTTCCAG 387
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TGGGATATATATGTCACACCTTAAGTACGCTCCACCTTCCCGAGGCGGAGGGGCGAG 390
QY 388 AATGATTTGGATTTGATTTGGAATCACAACACTCACACTCCCAAAAGAGTGGCATTTAG 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CGGTGATCTGTTTCACTTGGAGCCACCCCTTAAGTCCAGAGACTGGAGGCTTGAGC 450
QY 448 CACTGTTTAACCTGACTGCTGACTTACCGCGCTGATTCAGATATCCAAAGTGCCTTAAGC 507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AGATATCTCAATCTCACACTGTCATGCTTACCGAGAGACTCCGACATCTTCAAGCCCTTACGGC 510
QY 508 TTCTTGACGCTAAGCACAATCCCTTCGT-----GTTGAAGTGGCAAGCAAGAG 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 TGGCTGGAGCCGCTGTCCGCGCCAGCCTGCCACCCACCTCAACCTCTCGGCGCAAGACC 570
QY 559 AATTTGTTGCTGGTGTGAGTATGTAAGTCAAGCCTGAGCATGCGAGAGTCAAGTATTAC 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GAGCTGTGGCTGGCGGCTGTCTCAACTGGAAGCCGAGCTCAGCCAGGCTGCGCTACTAC 630
QY 619 AATGACTTAAGCAAAAGCATTTGAATTCATACCTACGAGGCGAAGCATTTGGAGATATGTC 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAGAGCTGCAAGGCTCATCTCAAGGTGAGCTGACGAGTACGAGAGCTCCACAAAGCCCTGCGCC 690
QY 679 AATGATTAATAATTTGATTTCTTCACTACATATGCTGTTAAATTTTATCTTCTTGAAGAT 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 AAGGGAGCA--TGATGAGAGAGCTGTCCCGGTACAAAGTCTACCTGCGCTTCGAGAAC 747
QY 739 TCAATCCACAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 TCTTTCGACCCCGACTACATCACTCAAGAGAGCTGTGAGAGAGCCCTGAGGCTGAGGCC 807
QY 796 GTAAGCTGTTGTTCTGTGGACATCTAGGAAACTATGAGATTAATTAATTAATTAATTAAT 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GTGCCGCTGTGCTGTGGGCCCCAGCAGCAAGCAACTACGAGAGGTCTGCGACCCGAGCGCC 867
QY 856 TTATTCATGTGGAATATATATACCTCCAGAGAGCTAGCAAAAGTATCTGAAGGAAGTC 915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 868 TTCATTCACGCTGAGACTTCCAGAGCCCAAGAGACGCGCCGCTACCTGAGAGGCTG 927

QY 916 GACAAAACATAAGTACTGTTAGTACTTACTAGTGAAG 960

Db 928 GACAAGGACGACGCGCGCTACTGAGCTACTTTCCTGCGCGGAG 972

RESULT 10

US-08-525-058A-12

Sequence 12 Application US/08525058A

Patent No. 5770420

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/525,058A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1086 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-525-058A-12

Query Match 10.4%; Score 112.2; DB 1; Length 1086;

Best Local Similarity 51.6%; Pred. No. 6e-24;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 271 TCCCATTCACAGGACGCTTCACTGTACAAACAATCCCATGCACTTCATCATCAG 330

Db 271 TCCCATTCACAGGACGCTTCACTGTACAAACAATCCCATGCACTTCATCATCAG 330

QY 331 CGAGCATCAGTTGGG---TCTGACAATTACTCTGAGCAAGTACGCCACCTTCAG 387

Db 331 TGGGATATCATGTCCAACTTAAGTACAGCCCTCCACCTTCCCGAGGCGAGGAG 390

QY 388 AATGATTTGGATGATTTGGATTCACCAACTCAGACTCCCAAAAGATGGCATTTAG 447

Db 391 CGCTGATCTGTTCACTGAGCCACCCCTTAAGTACGTCAGCACTTGAAGCCCTGAG 450

QY 448 CACTTGTAACTGACTGCTGACTTACCGCGGTGATTCAGATTCACCAAGTCTTATGC 507

Db 451 AGATACCTTCAATCTCACCATTCCACGAGCAGCTCCGACATCTTACGCGCTACGAG 510

QY 508 TTCTTGACGCTAAGCACAATCCCTTCGT-----GTTTGAAGTCCCAAGCAAGAG 558

Db 511 TGGCTGAGCGCTGTGCTCCGCGCAAGCTTCCACCCACCTCAACCTTCGCGCAAGACC 570

QY 559 AATATGCTGTGCTGGGTTGTGATTAATCTGAACCTTGAGATCCAGATCAATATTC 618

Db 571 GAGCTGTGCTGGCTGGGTGTCTCAATGGAACCGGACTACACGAGGTGGCTACTAC 630

QY 619 AATGAGCTAAGCAAAACATTTGAATCCATCTACGCGGCAACATTTGGAATATGTC 678

Db 631 CAGAGCTGACGCTCATCTCAAGTGAAGTGAAGTGAAGAGCTCCCAAGCCCTGCGCC 690

QY 679 AATGATAAAAATTTGATTCCTCATATCTGCTTGTAAATTTATCTTCTTGAAAT 738

Db 691 AAGGGAGCA---TGAAGGAGAGCTGTCCCGGTACAGTTCTACCTTGCTTCGAGAAC 747

QY 739 TCAATCCACAAGATTACATCAGCAAAAGCTAT--ACATGCTTTTCTGCTGCTCT 795

Db 748 TCTTTGACCCCGACTACATCAGCAAGAGCTGTGAGAGAGCCCTGGAGGCTGGGCC 807

QY 796 GTACCTGTGCTGAGCATCTAGGAAAACATATGAGAAATATTTCCAGCAGATTCA 855

Db 808 GTGCGCTGTGCTGGCGCCCAAGACCACTACAGAGGTTCCTGCCACCGCGCC 867

QY 856 TTCTATCATGTGGAAGTTATTAATCTCCAGAGCTGACCAAGATCTGAAGAGATC 915

Db 868 TTATTCACGCTGACGACTCCAGAGCCCAAGAGACTTGGCCGTAAGCTGAGAGAGCTG 927

QY 916 GACAAAACATAAGTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 960

Db 928 GACAAGGACGACGCGCGCTACTGAGCTACTTTCCTGCGCGGAG 972

RESULT 11

US-08-696-731-12

Sequence 12 Application US/08696731

Patent No. 595347

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/696,731

FILING DATE: 14-AUG-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/393,246

FILING DATE:

APPLICATION NUMBER: US 08/220,433

FILING DATE: 30-MAR-1994

APPLICATION NUMBER: US 07/914,281

FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

TELEX: 248855 OPAT DR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-696-731-12

Query Match 10.4%; Score 112.2; DB 2; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 66-24;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

271 TGGCATCTCACAAGGACGGCTTCACTGACACAAATCCCATGCTTGTGATCATCAG 330
 271 TGGCATCTCACAAGGACGGCTTCACTGACACAAATCCCATGCTTGTGATCATCAG 330
 331 CGAGCATCAGTTGGGA---TCGACAAATTTACCTCAGCAGCTAGCCACCTTCCAG 387
 331 TGGGATATCATGTCCAAACCTTAAGTACGCTCCACCTTCCCGGAGGCGCGAGGCGAG 390
 388 AATGATTTGGATGATTTGGATGACCACTACACACTCCCAAAAAGATGGCATTTAG 447
 391 CGCTGATCTGCTTCACTTGGAGCCACCCCTTAAGTCCAGCAGCCTGGAAGCCCTGGAC 450
 448 CACTGTTTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 507
 451 AGATCTTCAATCTCACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 510
 508 TTTCTGACGGTAAGCAAAATCCCTTCGT-----GTTGAAGTGGCAAGCAAGAG 558
 511 TGGCTGAGACCGCTGCTGCGGACCGCTGCGGACCGCTCAACCTCTGCGGCAAGACC 570
 559 AATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
 571 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
 619 AATGACTTAAGCAAAAGCAATTTGAAATCCATTAACGAGGCAAGCAATTTGAGAAATATGTC 678
 631 CAGAGCTCAGAGCTCATCTCAAGGTGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 690
 679 AATGATTAATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 738
 691 AAGGGAGCA---TGATGAGAGCGCTCCCGTACAGATTTCTTCTTCTTCTTCTTCTTCTTCT 747
 739 TCAATCTCAAGAGATTAATCTTCAAGGAAAGCTAT---ACAATGCTTTCTTCTTCTTCTTCT 795
 748 TCTCTGACACCCCGACTACATCAACGAGAAAGCTGTGAGGAGAGGAGGAGGAGGAGGAGGAG 807
 796 GAACTGTTGTTCTGAGGACATCTAGGAAACCTATGAGAAATTAATTTCTTCTTCTTCTTCT 855
 808 TGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
 856 TTTCTTCAATCTGAGAGATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 915
 868 TTTCTTCAATCTGAGAGATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 927
 916 GAAAAAACAATAAGTTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
 928 GAAAGAGACAGCCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 972

RESULT 12
 US-09-042-531-12

Sequence 12, Application US/09042531

Patent No. 6268193

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLION, SPIYAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,531

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/393,246

FILING DATE:

APPLICATION NUMBER: US 08/220,433

FILING DATE: 30-MAR-1994

APPLICATION NUMBER: US 07/914,281

FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1086 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

US-09-042-531-12

Query Match 10.4%; Score 112.2; DB 4; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 66-24;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

271 TGGCATCTCACAAGGACGGCTTCACTGACACAAATCCCATGCTTGTGATCATCAG 330
 271 TGGCATCTCACAAGGACGGCTTCACTGACACAAATCCCATGCTTGTGATCATCAG 330
 331 CGAGCATCAGTTGGGA---TCGACAAATTTACCTCAGCAGCTAGCCACCTTCCAG 387
 331 TGGGATATCATGTCCAAACCTTAAGTACGCTCCACCTTCCCGGAGGCGCGAGGCGAG 390
 388 AATGATTTGGATGATTTGGATGACCACTACACACTCCCAAAAAGATGGCATTTAG 447
 391 CGCTGATCTGCTTCACTTGGAGCCACCCCTTAAGTCCAGCAGCCTGGAAGCCCTGGAC 450
 448 CACTGTTTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 507
 451 AGATCTTCAATCTCACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 510
 508 TTTCTGACGGTAAGCAAAATCCCTTCGT-----GTTGAAGTGGCAAGCAAGAG 558
 511 TGGCTGAGACCGCTGCTGCGGACCGCTGCGGACCGCTCAACCTCTGCGGCAAGACC 570
 559 AATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
 571 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
 619 AATGACTTAAGCAAAAGCAATTTGAAATCCATTAACGAGGCAAGCAATTTGAGAAATATGTC 678
 631 CAGAGCTCAGAGCTCATCTCAAGGTGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 690

QY 679 AATGATATAAATTTGATTCCTACCAATATCTGCTGTAATTTTATCTTCCCTTTGAAT 738
DB 691 AAGGGAGCCA--TGATGAGAGCGCTGTCCGGTACCAAGTTTACCTGGCTTCGAGAAC 747
QY 739 TCAATCCACAAGATATACATCAGCAAGAAAGCTAT--ACAATGCTTTTCTGCTGCTCT 795
DB 748 TCTTTCACACCCCACTACATACACAGAGAGCTGTGAGAGAACCCCTGAGAGCCCTGGGCC 807
QY 796 GTACCTGTGTCTGTGGACCATCTAGGGAATACTATAGAAATTAATTCAGCAGATTCA 855
DB 808 GTCCCGGTGTGTGGGCCCCCAAGAGCACTACAGAGAGTCTTCCGCAACCCGAGGCC 867
QY 856 TTGATTCATGTGGAATTAATTAATCTCCAGAGTACCAAGATATCTGAAGGAAGTC 915
DB 868 TTGATTCAGTGTGAGAGCTTCCAGAGAGCCCAAGAGACTTGGCCCGGTACCTGAGAGCTC 927
QY 916 GACAAAACAATAAGTTATTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAAT 960
DB 928 GACAAAGACCAAGCCCGCTACCTGAGACTTTCGCTGGCGGAG 972

RESULT 13
US-09-063-237-2
Sequence 2, Application US/09063237
Patent No. 6124267

GENERAL INFORMATION:
APPLICANT: McEver, Rodger P. D.
APPLICANT: Cummings, Richard D.
TITLE OF INVENTION: O-glycan Inhibitors of Selectin Mediated
TITLE OF INVENTION: Inflammation Derived from PSGL-1
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30306-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/063,237
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/649,802
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMR#110CIP7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

Query Match 10.4%; Score 112.2; DB 3; Length 2042;
Best Local Similarity 51.6%; Pred. No. 8,2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;
QY 271 TGCATCTCAACAAGCAGCTTCACTGTACAAACAATCCCATGCACTTCTGATCCATCAC 330

DB 343 TGCACATCACTGACCCGCAAGAGGTATCCCAAGGAGAGACAGGTATCTGACACAG 402
QY 331 CGAGACATGATGTTGGA--TCTGACAAATTTACCTACAGCAAGCTAGGCAACCTTCAG 387
DB 403 TGGATATATATCTCCAACTTAAGTACAGCTTCCACCTTCCCGAGGCGGAGGAG 462
QY 388 AATGATTTGATGATGATTTGGAATCAGCACTCAACCTCCCAAAAGAGTGCAATGAG 447
DB 463 CCGTATGTTGTTCAACTTGGAGCACCCTTAATCTGACACCTGGAAGCCCTGAGC 522
QY 448 CACTGTTTAACTGACTGCTGACTACAGCGGCTGATGATGATGATGATGATGATGATG 507
DB 523 AGATACCTCATCTCAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 582
QY 508 TTCTTACCGGTAAAGCAAAATCCCTTCG-----GTTTGAAGTGCACAAAGAG 558
DB 583 TGGCTGAGACCGGTGTGCGGCGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 642
QY 559 AATTTGTTGTGCTGGTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 618
DB 643 GAGCTGTGTGCTGGCGGTGTGCAACTGGAAGCGGACTCAAGCGGTGCGCTACTAC 702
QY 619 AATGACTTAAGCAAAAGCAATTAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 678
DB 703 CAGAGCTGACAGCTCATCTCAAGTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 762
QY 679 AATGATATAAATTTGATTCCTACCAATATCTGCTGTAATTTTATCTTCTTTGAAT 738
DB 763 AAGGGAGCCA--TGATGAGAGCGCTGTCCGGTACCAAGTTTACTTGGCTTCGAGAAC 819
QY 739 TCAATCCACAAGATATACATCAGCAAGAAAGCTAT--ACAATGCTTTTCTGCTGCTCT 795
DB 820 TCTTTCACACCCGACTACATACACCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAG 879
QY 796 GTACCTGTGTCTGTGGACCATCTAGGGAATACTATAGAAATTAATTCAGCAGATTCA 855
DB 880 GTCCCGGTGTGTGGGCCCCCAAGAGCACTACAGAGAGTTCCTGCAACCCGAGGCC 939
QY 856 TTGATTCATGTGGAATTAATTAATCTTCCAGTACGTACCAAGATATCTGAAGAGTCT 915
DB 940 TTGATTCAGTGTGAGAGCTTCCAGAGAGCCCAAGAGAGCTTGGCCGCTTACCTGAGAGCTG 999
QY 916 GACAAAACAATAAGTTATTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAAT 960
DB 1000 GACAAAGACCAAGCCCGCTACCTGAGACTTTCGCTGGCGGAG 1044

RESULT 14
US-07-914-281-1
Sequence 1, Application US/07914281
Patent No. 5324663
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Tue Oct 8 10:18:27 2002

us-09-744-748-4_copy_11289_2368.rni

Page 12

```
Db 583 TGGCTGAGCCGCTGTCGCGCCGCTGCCCCACCCACCGCTCAACCTCTCGGCCAAGACC 642
QY 559 AATTTGTTGTGCTGGTTGTGAGTAAGTGAACCTGAGCATGCCAGAGTCAAGTAATAC 618
Db 643 GAGCTGGTGGCTGGGGGCTGCTCAACTGGAACCGGACTCAGCCAGGGTGGCTACTAC 702
QY 619 AATGAGCTAAGCAAAAGCATTTGAATCCATACCTACGGGCAAGCATTTGGAATATGTC 678
Db 703 CAGAGCTGCAAGCTATCTCAAGTGGAGCTGTACGGAGCTGCCACAAAGCCCCCTGCC 762
QY 679 AATGATTAATAATTTGATTCCTACCATATCTGTTGTAATTTATCTTCTTGAAT 738
Db 763 AAGGGGACCA---TGATGGAGAGCTGTCCCGGTACAACTTCTGAGCTTGAAGAC 819
QY 739 TCAATGCACAGAGATTACATGACGAAAGCTAT---ACAATGCTTTCTGGCTGCTCT 795
Db 820 TCCTTGACCCCGCTACTACATCACGAGAGAGCTGTGGAGAGAGCCCTGGAGGCTGGGCC 879
QY 796 GTACCTGTGTTGTTGGACCATCTAGGAAACTATGAGATTATATCCAGCAGATTCA 855
Db 880 GTGCCCGTGGCTGGGCCCCAGCAGAGCAACTACGAGAGGTTCCCTGCCACCCGACGCC 939
QY 856 TTCATTCAATGTGAGAGATTATTAACCTCTCCAGTAGCTAGCAAAAGTATCTGAAGAGTC 915
Db 940 TTCATCCAGCTGAGACACTTCCAGAGCCCCCAGAGACTGGCCCGGTACTCTGAGAGCTG 999
QY 916 GACAAAACAAATAAGTTACTTACTTACTTACTTACTGAGAGAG 960
Db 1000 GACAGAGACCAAGCCCGCTACTGAGCTACTTTCGCTGGGGGAG 1044
```

Search completed: October 6, 2002, 07:21:37
Job time : 29.0361 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 5, 2002, 22:39:45 ; Search time 1146.59 Seconds
(without alignments)
16699.717 Million cell updates/sec

Title: US-09-744-748-4_COPY_1454_2368

Perfect score: 915
Sequence: 1 accaaactgattatttaa.....agaaatgttttggaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	915	100.0	3019	9	AB023021	Homo sapi
2	913.4	99.8	80247	9	AL512406	AL512406 Human DNA
3	910.2	99.5	2501	9	HS238701	AJ238701 Homo sapi
4	792	86.6	1128	10	AF345993	AF345993 Rattus no
5	791.8	86.5	2139	10	AB015426	AB015426 Mus muscu
6	790.4	86.4	2156	10	AB049819	AB049819 Rattus no
7	783.8	85.7	1705	10	AF230460	AF230460 Cricetulu
8	666	72.8	1707	5	AB035905	AB035905 Gallus ga
9	581.2	63.5	1080	5	AB035905	AB035905 Xenopus l
10	580.2	63.4	71396	2	AC100674	AC100674 Mus muscu
11	277.8	30.4	1567	5	AB023627	AB023627 Dario rer
12	234.6	25.6	1832	5	AB023628	AB023628 Dario rer
13	222.6	24.3	1092	5	AB035907	AB035907 Xenopus l
14	220.2	24.1	1264	5	AF288369	AF288369 Gallus ga
15	176.2	19.3	2657	5	GGU73678	U73678 Gallus gall
16	145.2	15.9	1429	10	CGU78737	U78737 Cricetulus
17	145.2	15.9	1462	10	AF090450	AF090450 Cricetulu
18	143.6	15.7	1718	10	AF090449	AF090449 Cricetulu
19	137.2	15.0	1044	5	AB035908	AB035908 Gallus ga
20	132	14.4	2293	4	BT134FTG	X87810 B. taurus al
21	132	14.4	12607	4	BT1332772	AJ132772 Bos tauru
22	128	14.0	71396	2	AC100674	AC100674 Mus muscu
23	127.6	13.9	1158	10	AB039126	AB039126 Mus muscu
24	127	13.9	1487	3	AF016899	AF016899 Schistos
25	127	13.9	1814	6	AR028686	AR028686 Sequence
26	127	13.9	3594	10	MMU45980	U45980 Mus musculu
27	127	13.9	117993	2	AC091284	AC091284 Mus muscu
28	125.6	13.7	1613	4	BT1332776	AJ132776 Bos tauru
29	125.6	13.7	1619	4	BT1332775	AJ132775 Bos tauru
30	125.6	13.7	1644	4	BT1332774	AJ132774 Bos tauru
31	125.6	13.7	1873	9	AF345881	AJ132773 Bos tauru
32	124.6	13.6	1119	9	AF345881	AF345881 Macaca mu
33	122.8	13.4	1170	10	AB039127	AB039127 Mus muscu
34	122.8	13.4	1170	10	AB039131	AB039131 Mus muscu
35	122.8	13.4	1615	10	MMU33457	U33457 Mus musculu
36	122	13.2	1113	9	AF345884	AF345884 Saimiri s
37	121.2	13.2	1170	10	AB039124	AB039124 Mus muscu
38	121.2	13.2	1170	10	AB039125	AB039125 Mus muscu
39	121.2	13.2	1170	10	AB039128	AB039128 Mus muscu
40	121.2	13.2	1170	10	AB039129	AB039129 Mus muscu
41	121.2	13.2	1170	10	AB039130	AB039130 Mus muscu
42	121.2	13.2	1170	10	AB039132	AB039132 Mus muscu
43	119.8	13.1	1119	9	AF345883	AF345883 Macaca mu
44	118.4	12.9	1170	10	AB039133	AB039133 Mus spicl
45	118.2	12.9	1080	9	AB051859	AB051859 Pongo pyg

ALIGNMENTS

RESULT 1	AB023021	3019 bp	linear	PRI 23-JUN-1999
LOCUS	Homo sapiens FUT9 mRNA for alpha-1,3-fucosyltransferase IX,			
DEFINITION	complete cds.			
ACCESSION	AB023021			
VERSION	AB023021.1 GI:5139692			
KEYWORDS	alpha-1,3-fucosyltransferase IX; FUT9.			
SOURCE	Homo sapiens CDNA to mRNA.			
ORGANISM	Homo sapiens			
REFERENCE	1 (sites)			
AUTHORS	Kaneko,M., Kudo,T., Iwasaki,H., Ikehara,Y., Nishihara,S., Nakagawa,S., Sasaki,K., Shilina,T., Inoko,H., Saitou,N. and Narimatsu,H.			
TITLE	Alpha 1,3-fucosyltransferase IX (Fuc-TIX) is very highly conserved between human and mouse; molecular cloning, characterization and tissue distribution of human Fuc-TIX			
JOURNAL	FEBS Lett. 453, 237-242 (1999)			
REFERENCE	2 (bases 1 to 3019)			
AUTHORS	Kaneko,M., Kudo,T. and Narimatsu,H.			
TITLE	Direct Submission			

JOURNAL Submitted (29-JAN-1999) Mika Kaneko, Institute of Life Science,
Soka University, Division of Cell Biology, Soka University, 1-236
Tangai-cho, Hachioji, Tokyo 192-8577, Japan
(E-mail:mkanek@soka.ac.jp, tel:81-426-91-2495(ex.5132),
Fax:81-426-91-9315)

FEATURES
Location/Qualifiers
source 1..3019
/organism="Homo sapiens"
/db_xref="taxon:9606"
295..1374
/gene="FUT9"
295..1374
/gene="FUT9"
/codon_start=1
/product="alpha-L,3-fucosyltransferase IX"
/protein_id="BAA81685.1"
/db_xref="GI:5139693"
/translation="MTSTSKGILRPILVCIITIGCFMACILITIKPTNSMTSPESA
SSVLKMKPFSTKDYENETILWMPFGOTPLTSCQAFKSGIHEHNTITR
SHAVLIHRDLSMDLTNPOARPPROKIMTNESPHTQKSGIHEHNTITR
DSDIQPGPLEYSTNPFVPSKERVGVSNMNPERRAVTINLSKSIHTY
COAREYVNDKMLITISACFTLSFENSIRKDYITTELINAFIAGSPVPLGSRN
YENIIPADSFHVEDNPSSELAKILKEVDKNNKLYISYFMRADPTVNLPRWESHA
CLACHRVNRHOKISVGNLERKEMW"

BASE COUNT 937 a 594 c 541 g 947 t

ORIGIN

Query Match 100.0%; Score 915; DB 9; Length 3019;
Best Local Similarity 100.0%; Pred. No. 8.6e-231;
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ACCAAAGTATTTATTTATGAACTACTATCTGTGTGGGTGGCCATTTGGGAG 60
    |||||
DB 460 ACCAAAGTATTTATTTATGAACTACTATCTGTGTGGGTGGCCATTTGGGAG 519
    |||||
OY 61 ACCTTGACCTTATCCCGCAGCAATGTCACATCCAGATGCGATCTCACAG 120
    |||||
DB 520 ACCTTGACCTTATCCCGCAGCAATGTCACATCCAGATGCGATCTCACAG 579
    |||||
OY 121 GACCGTCTAGTACAAACATCCCATGAGTTCATCCATCCAGAGACAGTGTGG 180
    |||||
DB 580 GACCGTCTAGTACAAACATCCCATGAGTTCATCCATCCAGAGACAGTGTGG 639
    |||||
OY 181 GATCTACAAATTTACTAGCAAGTATGCCACCTTCAGAAATGGATTGGATG 240
    |||||
DB 640 GATCTACAAATTTACTAGCAAGTATGCCACCTTCAGAAATGGATTGGATG 699
    |||||
OY 241 TTGGAAATCAACATCCATCCCAAAAGATGGCATTTGACACTTGTAACTG 300
    |||||
DB 700 TTGGAAATCAACATCCATCCCAAAAGATGGCATTTGACACTTGTAACTG 759
    |||||
OY 301 CTGACTTACCGCGGATTCAGATATCAAGTCCCTTATGCTCTTCAAGGTAGCA 360
    |||||
DB 760 CTGACTTACCGCGGATTCAGATATCAAGTCCCTTATGCTCTTCAAGGTAGCA 819
    |||||
OY 361 AATCCCTGTGTGTTGAATGCCAAGCAAGAAATGGTGTGCTGGTGTGATG 420
    |||||
DB 820 AATCCCTGTGTGTTGAATGCCAAGCAAGAAATGGTGTGCTGGTGTGATG 879
    |||||
OY 421 TTGAACCCGAGATGCGAGTCAAGTTTACATGAGCTTAGCAAAAGATTTAAATC 480
    |||||
DB 880 TTGAACCCGAGATGCGAGTCAAGTTTACATGAGCTTAGCAAAAGATTTAAATC 939
    |||||
OY 481 CATACCTAGGGGAGCAATTTGGAGAAATGTCAATGATATAAATTTGATCTTACATA 540
    |||||
DB 940 CATACCTAGGGGAGCAATTTGGAGAAATGTCAATGATATAAATTTGATCTTACATA 999
    |||||
OY 541 TCTGCTTGAATTTTATCTTCTTGAATTTCAATCCACAGATTTACATCAGGAA 600
    |||||
DB 1000 TCTGCTTGAATTTTATCTTCTTGAATTTCAATCCACAGATTTACATCAGGAA 1059
    |||||
OY 601 AAGCTATACATGCTTTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 660
    |||||

```

Db 1060 AAGTATATCAATGCTTTTCTGTGGCTGCTGTACTGTGTGTCTGTGGACCATCTAGGAA 1119

OY 661 AACTATGAGATATATATTCACAGAGATTCATTCATTCATGTGGAAGATTTAACTCTCC 720

Db 1120 AACTATGAGATATATATTCACAGAGATTCATTCATTCATGTGGAAGATTTAACTCTCC 1179

OY 721 AGTAGCTACCAAGATATCTGAAGAGTGCACAAAACAAATTAATTAATCTAGTTAGTAC 780

Db 1180 AGTAGCTACCAAGATATCTGAAGAGTGCACAAAACAAATTAATTAATCTAGTTAGTAC 1239

OY 781 TTTAAGTGAAGAGATTTACACTGTAATCTCCACGATTTTGGAAATCATGATGATG 840

Db 1240 TTTAAGTGAAGAGATTTACACTGTAATCTCCACGATTTTGGAAATCATGATGATG 1299

OY 841 TTGGCTTGCATGATGGAAGAGCAATCAAGATTAATGCTGCTGTGTAATTAAGAA 900

Db 1300 TTGGCTTGCATGATGGAAGAGCAATCAAGATTAATGCTGCTGTGTAATTAAGAA 1359

OY 901 TGGTTTGGAAATTA 915

Db 1360 TGGTTTGGAAATTA 1374

RESULT 2

AL512406

LOCUS

DEFINITION Human DNA sequence from clone RP11-504J9 on chromosome 6, complete

ACCESSION AL512406

VERSION AL512406.14 GI:13897154

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 80247)

AUTHORS Bales K.

JOURNAL Direct Submission

COMMENT

Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

request: clone request@sanger.ac.uk

On Apr 30, 2001 this sequence version replaced gi:13396709.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sg: SWISSPROT; Tr: TrEMBL; Wp: WormERP; Information on the WormERP database can be found at

<http://www.sanger.ac.uk/Projects/Celegans/wormerp> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>

RP11-504J9 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/dacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-504J9. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP3-381A2 is at 80148 in this sequence. The true right end of clone RP11-77B15 is at 100 in this sequence.

```
FEATURES
source
Location/Qualifiers
1..60247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone_id="RP11-504J9"
/clone_lib="RPCT-11.2"
184..1422
repeat_region repeat: matches 875..2419 of consensus"
1744..1834
repeat_region repeat: matches 166..258 of consensus"
2950..2977
repeat_region repeat: matches 2 mer ca 100% conserved"
3411..3669
repeat_region repeat: matches 93..364 of consensus"
4372..4704
repeat_region repeat: matches 5..365 of consensus"
complement(7821..8406)
misc_feature repeat: matches 504..542 of consensus"
8126..8184
repeat_region repeat: matches 272..332 of consensus"
8185..8696
repeat_region repeat: matches 1..513 of consensus"
8779..8817
repeat_region repeat: matches 504..542 of consensus"
8819..8974
repeat_region repeat: matches 6..162 of consensus"
complement(9443..9891)
misc_feature repeat: matches 6..162 of consensus"
complement(9452..9883)
misc_feature repeat: matches 92..205 of consensus"
9782..9894
repeat_region repeat: matches 135..262 of consensus"
10089..10212
repeat_region repeat: matches 10671
complement(10093..10671)
misc_feature repeat: matches 203..262 of consensus"
10223..10282
repeat_region repeat: matches 203..262 of consensus"
complement(10315..10631)
misc_feature repeat: matches 203..262 of consensus"
10805..11186
repeat_region repeat: matches 2546..2750 of consensus"
11657..11844
repeat_region repeat: matches 2546..2750 of consensus"
complement(11659..12084)
misc_feature repeat: matches 2546..2750 of consensus"
complement(11674..11987)
misc_feature repeat: matches 2546..2750 of consensus"
12083..12588
repeat_region repeat: matches 2546..2750 of consensus"
12596..12810
repeat_region repeat: matches 2546..2750 of consensus"
12920..12986
repeat_region repeat: matches 2546..2750 of consensus"
12987..13169
repeat_region repeat: matches 2546..2750 of consensus"
13168..13280
repeat_region repeat: matches 2546..2750 of consensus"
13851..17218
repeat_region repeat: matches 2546..2750 of consensus"
18903..18953
repeat_region repeat: matches 2546..2750 of consensus"
18984..19433
repeat_region repeat: matches 2546..2750 of consensus"
complement(19108..19690)
misc_feature repeat: matches 2546..2750 of consensus"
20683..20801
repeat_region repeat: matches 2546..2750 of consensus"
20802..20862
repeat_region repeat: matches 2546..2750 of consensus"
20912..21033
repeat_region repeat: matches 2546..2750 of consensus"
21047..21128
repeat_region repeat: matches 2546..2750 of consensus"
21676..21859
repeat_region repeat: matches 2546..2750 of consensus"
21890..23612
repeat_region repeat: matches 2546..2750 of consensus"
23601..26609
repeat_region repeat: matches 2546..2750 of consensus"
26620..26789
repeat_region repeat: matches 2546..2750 of consensus"
28521..28627
repeat_region repeat: matches 2546..2750 of consensus"
28636..28837
repeat_region repeat: matches 2546..2750 of consensus"
28838..29142
repeat_region repeat: matches 2546..2750 of consensus"
29143..29484
repeat_region repeat: matches 2546..2750 of consensus"
30529..31344
repeat_region repeat: matches 2546..2750 of consensus"
32234..32340
repeat_region repeat: matches 2546..2750 of consensus"
32399..32578
repeat_region repeat: matches 2546..2750 of consensus"
32816..32979
repeat_region repeat: matches 2546..2750 of consensus"
33312..33596
repeat_region repeat: matches 2546..2750 of consensus"
34856..34915
repeat_region repeat: matches 2546..2750 of consensus"
complement(34987..35447)
misc_feature repeat: matches 2546..2750 of consensus"
match: STS: Em:AQ321598
complement(36488..36936)
misc_feature repeat: matches 2546..2750 of consensus"
complement(36654..37090)
misc_feature repeat: matches 2546..2750 of consensus"
37102..37591
repeat_region repeat: matches 2546..2750 of consensus"
complement(37581..38221)
misc_feature repeat: matches 2546..2750 of consensus"
38218..38732
repeat_region repeat: matches 2546..2750 of consensus"
38861..39164
repeat_region repeat: matches 2546..2750 of consensus"
39612..39653
repeat_region repeat: matches 2546..2750 of consensus"
40854..40909
repeat_region repeat: matches 2546..2750 of consensus"
42169..42270
repeat_region repeat: matches 2546..2750 of consensus"
42411..42548
repeat_region repeat: matches 2546..2750 of consensus"
42780..42938
repeat_region repeat: matches 2546..2750 of consensus"
43229..43326
repeat_region repeat: matches 2546..2750 of consensus"
45383..45428
repeat_region repeat: matches 2546..2750 of consensus"
complement(47284..47971)
misc_feature repeat: matches 2546..2750 of consensus"
match: STS: Em:AQ269059
complement(47537..48011)
misc_feature repeat: matches 2546..2750 of consensus"
complement(47723..48014)
```



```
Db 330 ACCCTTGACCTTACATCCATCGCCAGCAAGTGTCAACATCCAGGATCCATCTCACAAAG 389
Oy 121 GACCGTTCACTGTACAAACAAATCCCATGAGTTCGTATCCATCCAGCAGACATCACTGG 180
Db 390 GACCGTTCACTGTACAAACAAATCCCATGAGTTCGTATCCATCCAGCAGACATCACTGG 449
Oy 181 GATCTGACAAATTTTACCTCAGCAAGCTAGCCACCTTCAGAAATGATTTGGATGAAT 240
Db 450 GATCTGACAAATTTTACCTCAGCAAGCTAGCCACCTTCAGAAATGATTTGGATGAAT 509
Oy 241 TTGGAAATCAACCACTCACACTCCCCAAAAGATGGATGGATGAGCAGCTTTTAACCTGACT 300
Db 510 TTGGAAATCAACCACTCACACTCCCCAAAAGATGGATGGATGAGCAGCTTTTAACCTGACT 569
Oy 301 CTGACTTACCGCGGTGATTCAGATATCCAAAGTCCCTTATGGCTTTCTTGACGGTAAACACA 360
Db 570 CTGACTTACCGCGGTGATTCAGATATCCAAAGTCCCTTATGGCTTTCTTGACGGTAAACACA 629
Oy 361 AATCCCTTCGTGTGAAAGTCCCAAGCAAGAAATTTGGTGTCTGGCTGTGAGTAAAC 420
Db 630 AATCCCTTCGTGTGAAAGTCCCAAGCAAGAAATTTGGTGTCTGGCTGTGAGTAAAC 689
Oy 421 TGGAAACCTGAGACATGCCAGAGTCAAGTATTTACATGAGCTTAAGCAAAAGCATTTGAATC 480
Db 690 TGGAAACCTGAGACATGCCAGAGTCAAGTATTTACATGAGCTTAAGCAAAAGCATTTGAATC 749
Oy 481 CATACCTACGGGCAAGCATTTTGAGAAATATGCAATGATAAATTTGATTCCTTACCATA 540
Db 750 CATACCTACGGGCAAGCATTTTGAGAAATATGCAATGATAAATTTGATTCCTTACCATA 809
Oy 541 TCTGCTTGTAAATTTTATCTTCTTCTTGAATAATTCATCCACAGAGATTACATCAGGAA 600
Db 810 TCTGCTTGTAAATTTTATCTTCTTCTTGAATAATTCATCCACAGAGATTACATCAGGAA 869
Oy 601 AAGCTATTAATATGCTTTTGTGGCTGCTGTACCTGTTGTTCTGGAGACATCTAGAGAA 660
Db 870 AAGCTATTAATATGCTTTTGTGGCTGCTGTACCTGTTGTTCTGGAGACATCTAGAGAA 929
Oy 661 AACTATGAAATTTATTTTCCAGACATTCATTCATGATGGAATATATACTCTCC 720
Db 930 AACTATGAAATTTATTTTCCAGACATTCATTCATGATGGAATATATACTCTCC 989
Oy 721 AGTGAAGTGAAGAGTATCTGAAGAGTGAAGAGTGAAGAGTGAAGTATTAAGTATTAAGT 780
Db 990 AGTGAAGTGAAGAGTATCTGAAGAGTGAAGAGTGAAGAGTGAAGTATTAAGTATTAAGT 1049
Oy 781 TTTTAACTGAGAGAGATTTTCACTGTAATCTTCCAGATTTTGGAGATCAGATGATGT 840
Db 1050 TTTTAACTGAGAGAGATTTTCACTGTAATCTTCCAGATTTTGGAGATCAGATGATGT 1109
Oy 841 TTGGCTTTGGATCATGTGAAAGAGCATCAAGATATTAAGTGTGGTATTTAGAGAA 900
Db 1110 TTGGCTTTGGATCATGTGAAAGAGCATCAAGATATTAAGTGTGGTATTTAGAGAA 1169
Oy 901 TGGTTTGGGAATTA 915
Db 1170 TGGTTTGGGAATTA 1184

RESULT 4
AF345993 1128 bp mRNA linear ROD 07-MAR-2001
LOCUS Rattus norvegicus alpha1,3-fucosyltransferase IX (FucT) mRNA,
DEFINITION complete cds.
ACCESSION AF345993
VERSION AF345993.1 GI:13242183
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
```

```
REFERENCE 1 (bases 1 to 1128)
AUTHORS Baboval,T., Henton,T., Kinnally,E. and Smith,F.I.
TITLE Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX)
and comparison of the expression of fuc-TIV and fuc-TIX genes
during rat postnatal cerebellum development
JOURNAL J. Neurosci. Res. 62 (2), 206-215 (2000)
PUBMED 11020213
REFERENCE 2 (bases 1 to 1128)
AUTHORS Smith,F.I. and Baboval,T.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver
Center, 200 Trapelo Rd., Waltham, MA, USA
FEATURES
source 1..1128
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1128
/gene="Fuc9"
49..1128
/gene="Fuc9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="AAK16591.1"
/db_xref="GI:13242184"
/translation="MTSTSKGILREPLIVCIILIGCFVACLIIYIKPTNSWFSWPESA
SSVLKMKNEFSKRDYFENETILVWVMEFGOTFEDSCAMENIGCHLIDRLSYNK
SHAVLIHRDLSMDLNLDOARPEPKWIMNLESPIHTPOKSGIEHLFTLYR
DSDIQYRGLFVTSPPFVPEPSKEKLVYVSWNEBHARVRYINLSSTIEHTY
GOAREIYNDKMLPFTSTCKRYLSFENSIRHDIYTERLYNAFLAGSVPAVLGPREN
YENIYPADSFLIHVEDFNPSELAFLKDKNNKLYLSPFWRKDFVNLPRFWSHA
CLACDHVRHDEYKSVGNLEKFWFM"
BASE COUNT 324 a 256 c 228 g 320 t
ORIGIN
Query Match 86.6%; Score 792; DB 10; Length 1128;
Best Local Similarity 91.8%; Pred. No. 2.5e-198;
Matches 837; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Oy 4 AAACGATTAATTTTAAATGAAATCTATCTGAGTGGCTGGCCATTTGGCGAGACC 63
Db 217 AAACGATTAATTTTAAATGAAATCTATCTGAGTGGCTGGCCATTTGGCGAGACC 276
Oy 64 TTTGACCTTACATCCCGCCAGCAAGTGTCAACATCCAGAGATGCCATCTCAACAGGAC 123
Db 277 TTTGACCTTACATCCCGCCAGCAAGATGTTCAACATCCAGAGATGCCATCTCAACAGGAC 336
Oy 124 CGTTACGTGACAAATTCATGCAAGTGTGATTCATACCGAGACATCAGTTGGAT 183
Db 337 CGCTCGTTGACAAATTCATGCAAGTGTGATTCATACCGAGACATCAGTTGGAT 396
Oy 184 CTGACAAATTTACCTCAGCAAGTGTGATTCATACCGAGACATCAGTTGGAT 243
Db 397 CTGACAAATTTACCTCAGCAAGTGTGATTCATACCGAGACATCAGTTGGAT 456
Oy 244 GATCACCACACTCACACTCCCAAAAGAGTGGCATGAGCAGCTGTTTAACCTGACTCG 303
Db 457 GATCACCACACTCACACTCCCAAAAGAGTGGCATGAGCAGCTGTTTAACCTGACTCG 516
Oy 304 ACTTACCGCGGTGATTCAGATATCCAAAGTCCCTTATGGCTTTGACGGTAAACAAAT 363
Db 517 ACTTACCGCGGTGATTCAGATATCCAAAGTCCCTTATGGCTTTGACGGTAAACAAAT 576
Oy 364 CCCTTGCTTTGGAAGTGGCAAGCAAGAAATGAGTGGTGGTGGTGGTGGTAACTG 423
Db 577 CCCTTGCTTTGGAAGTGGCAAGCAAGAAATGAGTGGTGGTGGTGGTGGTAACTG 636
Oy 424 AACCTGACATCCAGAGTCAAGTATTTCAATGAGCTAAGCAAAAGCATTTGAATTCAT 483
Db 637 AATCCCGACATCCAGAGTCAAGTATTTCAATGAGCTAAGCAAAAGCATTTGAATTCAT 696
Oy 484 AACTACGGGCAAGCATTTGGAATATTTGCAATGATAAAAATTTGATTCATACATATCT 543
```

Db	697	ACCAATGCGACACCAATTTGGAGAAATATGTGAATGATATAAAATCTGATTTCCCAACCAATATCT	756
QY	544	GCTTGTAAATTTTATCTTTCTTTGAAAATTCATCCACAAGATTTACATCAGCGAAAAG	603
Db	757	ACTGTAAATTTTATCTTTCTTTGAAAACCAATTCACAAAGATTTACATCAGCAAGAAAAG	816
QY	604	CTATACAAATGCTTTCTGGCTGGCTCTGTACTGTGTGCTGGAGACCATGTAGAGAAAAC	663
Db	817	CTCTACAAATCATTTTGTGGCTGGTCTGTGACTGTGTGCTGGAGTCTAGAGAAAAC	876
QY	664	TATGAGATTTATTTCCAGACATTCATTCATCACTGTGGAAGTTTAACTCTCCAGT	723
Db	877	TATGACAAATTTATTTCCAGCTATTCATTCATTCATGTGGAAGATTTTAACTCTCCAGT	936
QY	724	GAGCTACAAAGTATCTGAAAGAGTGCACAAAACAAATAAGTATTACCTTAACTTACTTT	783
Db	937	GAGTGTGCCAAAATCTGGAAGAGGTTCACAAAACATAAAGTATTCTTAACTTACTTCTC	996
QY	784	AACGTGAGAGAGATTTCACTGTAATCTCCAGATTTTGGCAATACATGATGATTTTGG	843
Db	997	AACGTGAGAGAGATTTTCACTGTAATCACTCCAGCGTTTGGGAATACATGATGATCTCG	1056
QY	844	GCTTGGATCATGTGAAAAAGGATCAAGATATTAAGTCTGTGTGTAATTTAGAGAAATGG	903
Db	1057	GCAATGCATCATGTATAAAAGGCATCAGAGATATTAAGTCTGTGTGCAATTTAGAGAAATGG	1116
QY	904	TTTTGGAATTTA 915	
Db	1117	TTTTGGAATTTA 1128	

LOCUS	2139 bp	mRNA	linear	ROD 23-JUN-1999
DEFINITION	AB015426	Mus musculus	Fut3 mRNA for alpha,3'-fucosyltransferase IX, complete	
ACCESSION	AB015426			
VERSION	AB015426.1	GI:3702718		
KEYWORDS	Fut3; alpha1,3-fucosyltransferase IX.			
SOURCE	Mus musculus cDNA			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (sites)			
TITLE	Kudo,T., Ikehara,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K. and Narimatsu,H.			
JOURNAL	Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mFuc-ITX, that synthesizes the Lewis x (CD15) epitope in brain and kidney			
MEDLINE	J. Biol. Chem. 273 (41), 26729-26738 (1998)			
REFERENCE	98434588			
AUTHORS	2 (bases 1 to 2139)			
TITLE	Kudo,T. and Narimatsu,H.			
JOURNAL	Direct Submission			
DEFINITION	Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science, Soka University, Division of Cell Biology, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan (E-mail:tkudo@et.soka.ac.jp, tel:+81-426-91-2495, Fax:+81-426-91-9315)			
FEATURES	Location/Qualifiers			

FEATURES	Location/Qualifiers
source	1..2139
	/organism="Mus musculus"
	/db_xref="taxon:10090"
gene	115..1194
	/gene="Puf9"
CDS	115..1194

```

/gene="Put9"
/codon_start=1
/product="alpha,3-fucosyltransferase IX"
/protein_id="AA3352.1"
/db_xref="GI:3702719"
/translation="MTSTSGILIRPLIYCIIIGCMACILYIPETSWFSPMESNA
SHAVLKKHNSFSTKDYNEETIIIVWVPGCFDILSCQAMNIOGCHLITDRLSYL
SVATLHKHRSIDSLNIDY,PDAAAPPPKMTKNTWNIISPTTPPKSGTEHLFILTIYYRR

```

[illegible]

Query Match	Similarity	Score	DR	Length
Best Local	85.5%	791.8	10	2139
Matches	91.6%	Pred. No. 2	198	
838	Conservative	0	Mismatches	77
			Indels	0
			Gaps	0
QY 1	ACCAAAACGATATTTTAATGAACATCTATCTGTGTGGGTGGGCCATTTGGGCAG	60		
Db 280	ACAAAAACGATATTTTAATGAACATCTATCTGTGTGGGTGGGCCATTTGGGCAG	339		
QY 61	ACCTTTGACCTTAACATCTCTCCCAACCAATGTTCAACATCCAAAGATGCCATCTCAACG	120		
Db 340	ACCTTTGACCTTAACATCTCTCCCAACCAATGTTCAACATCCAAAGATGCCATCTCAACG	399		
QY 121	GACCGTTCACTAGTACAAACAAATCCCATCAATTCATCATCCGAAACATCAAGTTGG	180		
Db 400	GACCGCATATTTGACAAACAAATCCCATCGGTCTGATTCACCATATGACATCAAGCTGG	459		
QY 181	GATGTGCAATTTTACCTCAGAACCTTAGGCGCACCTTCACGAAATATGAAATTTGATGAT	240		
Db 460	GATGTGCAATTTTACCTCAGAACCTTAGGCGCACCTTCACGAAATATGAAATTTGATGAT	519		
QY 241	TTGGAAATCAACCACTCAGCTCCCAAAAGATGGCATAGACATGTTTAACTGACT	300		
Db 520	TTAGAGTCAACCACTCAGCTCCCAAAAGATGGCATATGAACTTTGTTCACTGACT	579		
QY 301	CTGACTTACCGCCGTGATTCAGATATCCAAATGGCTTATGGCTTTTCAACGGTAAACCA	360		
Db 580	CTACTATTCGCGGTATTCAGATATCCAAATGGCTTATGGCTTTTCAACGGTAAACCA	639		
QY 361	AATCCCTTGCTGTTTGAAGTGGCAAGCAAGAAATTTGTTGCTGGGTTGTGATTAAC	420		
Db 640	AATCCCTTGCTGTTTGAAGTGGCAAGCAAGAAATTTGTTGCTGGGTTGTGATTAAC	699		
QY 421	TGGAACCTGAGCATCCAGAGTCAAGTATTACAAATGAGCTTAAGCAAAAGCATTTGAATC	480		
Db 700	TGGAACCTGAGCATCCAGAGTCAAGTATTACAAAGTCAAGAGTATTGAAATC	759		
QY 481	CATACCTAGGGGCAACATTTTGAGAAATATGTCAATATATTAATTTTGATTCCTACATA	540		
Db 760	CACACCTAGGGGCAACATTTTGAGAAATATGTCAAGTCAAGCATATTAATTTTGATTCCTACATA	819		
QY 541	TCGTGCTGTAATTTTATCTCTCTTGAATAATTCATTCACAAAGATTATACACAGAA	600		
Db 820	TCGTGCTGTAATTTTATCTCTCTTGAATAATTCATTCACAAAGATTATACACAGAA	879		
QY 601	AAGGTATACATCTTTCTGCTGGCTGTCTACTCTGTCTGGGACCATCTAGGGAA	660		
Db 880	AAGGTATACATCTTTCTGCTGGCTGTCTACTCTGTCTGGGACCATCTAGGGAA	939		
QY 661	AACATAGGAATATATATTCGACAAATTCATCTATCTATGGAATAATATATTAACCTCC	720		
Db 940	AACATAGGAATATATATTCGACAAATTCATCTATCTATGGAATAATATATTAACCTCC	999		
QY 721	AGTAGAGTACCAAGTATCTGGAAGAGTGCACAAAACAAATATGATTAATCTTAGTTAC	780		
Db 1000	AGTAGAGTACCAAAATATCTGGAAGAGTGCACAAAACAAATATGATTAATCTTAGTTAC	1059		
QY 781	TTTAACTGAGGAAGGATTTCACTCTTAATTTTCCACGATTTTGGGAATCCACATGCATGT	840		
Db 1060	TTTAACTGAGGAAGGATTTCACTCTTAATTTTCCACGATTTTGGGAATCCACATGCATGT	1119		
QY 841	TTGGCTTGCGATATATGAAAAGCATCAAAATATATGAAGCTCTTGATTAATTAAGAAA	900		
Db 1120	CTGGCATATGCGACATATATAAAGCATCAAAATATATGAAGCTCTTGATTAATTAAGAAA	1179		

QY 901 TGGTTTGAATTA 915
Db 1180 TGGTTTGAATTA 1194

RESULT 6
AB049819

LOCUS AB049819 2156 bp mRNA linear ROD 18-JUN-2002
DEFINITION Rattus norvegicus FUR9 mRNA for alpha1,3-fucosyltransferase IX,
complete cds.
AB049819
AB049819.1 GI:13591588

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
AUTHORS
TITLE
JOURNAL

1 (sites)
Shimoda,Y., Tajima,Y., Osanai,T., Katsune,A., Kohara,M., Kudo,T.,
Narimatsu,H., Takashima,N., Ishii,Y., Nakamura,S., Osunli,N. and
Sanai,Y.
Pax6 Controls the Expression of Lewis x Epitope in the Embryonic
Forebrain by Regulating alpha 1,3-Fucosyltransferase IX Expression
J. Biol. Chem. 277 (3), 2033-2039 (2002)
11675393
2 (bases 1 to 2156)
Sanai,Y.
Direct Submission
Submitted (11-OCT-2000) Yutaka Sanai, Tokyo Metropolitan Institute
of Medical Science, Department of Biochemical Cell Research;
Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113-8613, Japan
(E-mail:sanaie@rinsoken.or.jp, Tel:81-3-3823-2101(ex.5233),
Fax:81-3-3828-6663)
Location/Qualifiers
1. 2156
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
346..1425
/gene="FUT9"
346..1425
/gene="FUT9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="BAB40953.1"
/db_xref="GI:13591589"
/translation="MTSTSKGILRPFLVCIITIGCFVACLIYIKPTNSWVSPWESA
SVLKKMKPFSPKDYFNTETTLVWVWPGQFDLSCQAMNIOGCHITTPRSLYNK
SHAVLIHHHDSMDLNLDOARPPFOKIMNLSPHTPOKSGIENHFNLIYIRK
DSDIQVPGFLVSTSPVPEVPSKELVWVSWNPENHAKVYINELSKSIEIHTY
GOAFGEVYNDKMLIPTISTCKEYLSFENSHDYITTEKLYNAFLAGSYVAVGSPREN
YENIIPADSEFIVEDFNSPSELAKYLEVDKNNKLYLSYFNMRKDFVTLPRFWESHA
CLACDHVRHOKESVGNLEKWEFN"

BASE COUNT 605 a 470 c 437 g 644 t
ORIGIN

Query Match 86.4%; Score 790.4; DB 10; Length 2156;
Best Local Similarity 91.7%; Pred. No. 6.5e-198;
Matches 836; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 4 AAAACGATTTATTTAATGAACACTATTTCGTGGTGGTGGCCATTTGGGCGAGCC 63
Db 514 AAAACGATTTATTTAATGAACACTATTTCGTGGTGGTGGCCATTTGGGCGAGCC 573

QY 64 TTGACCTTACATCTGCGAAGCATGTTCACATCCAGGATGCATCTCACAGGAGC 123
Db 574 TTGACCTTACATCTGCGAAGCATGTTCACATCCAGGATGCATCTCACAGGAGC 633

QY 124 CGTTACTGTACAAACATCCATGAGTTCTGATCCATCACCAGACATCACTGGGAT 183
Db 634 CGTGTGTGTACAAACATCCATGAGTTCTGATCCATCACCAGACATCACTGGGAT 693

QY 184 CTGACAAATTTACCTGACGACAGCCACCTTCACAGAAATGATTTGGATGAATTTG 243

Db 694 CTGACTTAACCTTACCTGACGAGCCAGCCGCCCTTTCAGAAATGATTTGGATGAATTTG 753

QY 244 GAATCACCAACTACACACTCCCAAAAAGAGTGGATGAGACTTGTATTAACCTGACTG 303
Db 754 GAGTCAACCACTACACACCCCAAAAAGAGTGGATGAGACTTGTATTAACCTGACTG 813

QY 304 ACTTACCGCCGTGATTCAGATATTCACAGTCCCTTATGCTCTTGAACGGTAAACCAAT 363
Db 814 ACTTATGCGCGGTATTCAGATATTCACAGTCCCTTATGCTCTTGAACGGTAAACCAAG 873

QY 364 CCCTTGTGTTTGAAGTGGCAAGCAAGAAATTTGGTGTGCTGGGTGTGATTAACCTG 423
Db 874 CCCTTGTGTTTGAAGTGGCAAGCAAGAAATTTGGTGTGCTGGGTGTGATTAACCTG 933

QY 424 AACCTGAGACCCAGAGTATTCAGATATTCAGATTAAGACTTAAGCAAAAGATTAATCCAT 483
Db 934 AATCCGAGACATCCAGAGTATTCAGATTAAGACTTAAGCAAAAGATTAATCCAT 993

QY 484 ACCTAGCGGACCAATTTGGAGAAATATTCATATGATTAATTAATTTGATTCCTACCATATCT 543
Db 994 ACCTATGCGCAACATTTGGAGAAATATTCATATGATTAATTAATTTGATTCCTACCATATCT 1053

QY 544 GCTTGTAAATTTATCTTCTTCTTGAATTCATCAAGATTAACATCAAGATTAACATCAAG 603
Db 1054 ACTTGTAAATTTATCTTCTTCTTGAATTCATCAAGATTAACATCAAGATTAACATCAAG 1113

QY 604 CTATACATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
Db 1114 CTATACATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1173

QY 664 TATGAGAAATTAATTCACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 723
Db 1174 TATGAGAAATTAATTCACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1233

QY 724 GAGCTAGCAAGATTCATTCAGAGAGCGACAAAACAAATTAATTAATTAATTAATTAATTA 783
Db 1234 GAGTGTGCAAAATTAATTCAGAGAGCGTTCACAAAACAAATTAATTAATTAATTAATTA 1293

QY 784 AACTGGAGAGAGATTCATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 843
Db 1294 AACTGGAGAGAGATTCATTCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1353

QY 844 GCTTGGATCATGTGAAGAGCATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 903
Db 1354 GCAATGAGATCATGTGAAGAGCATCAAGAAATTAATTAATTAATTAATTAATTAATTA 1413

QY 904 TTTTGAATTTAA 915
Db 1414 TTTTGAATTTAA 1425

RESULT 7
AF230460

LOCUS AF230460 1705 bp mRNA linear ROD 12-JUL-2000
DEFINITION Citreolus griseus alpha(1,3)fucosyltransferase type IX (Fut9)
mRNA, complete cds.
AF230460
AF230460.1 GI:9049663

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Chinese hamster.
Citreolus griseus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE

1 (bases 1 to 1705)
Patnaik,S.K., Zhang,A., Shi,S. and Stanley,P.
Alpha(1,3)fucosyltransferases expressed by the gain-of-function
Chinese hamster ovary glycosylation mutants IEC12, IEC29, and IEC30
Arch. Biochem. Biophys. 375 (2), 322-332 (2000)
2016953
10700388
2 (bases 1 to 1705)

AUTHORS Patnaik, S.K., Shi, S. and Stanley, P.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of
Medicine, Chanin 516, 1300 Morris Park Avenue, New York, NY 10461,
USA

FEATURES

source location/Qualifiers
1. 1705
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
1. 1705
/gene="Fut9"
/note="cgFut9"
62. 1141
/gene="Fut9"
/note="Fuc-TIX"
/codon_start=1
/product="alpha(1,3)fucosyltransferase type IX"
/protein_id="AAF82412.1"
/db_xref="GI:9049664"
/translation="MSTSKGILRPFLIVCYTLACPMACILYIKPTSNVSPWESA
SAYLKKNFESTKDYENETTLVWMPGQFDTSCQAMPNIGCHLTDRLYRK
SHAVLHHRDIDMDLNLPOQARPPQKIMANLSPHTPOKSEIHLNLTLYRK
DSDIQYPIGFLVSTNPEYEVPSKEKLVCWVSNWNEHARVKYINELSLYIRY
GOAFGEYVNDKMLIPTISCKFLYSESKELVSIKDYITTELYNALAGSPVVLGSRN
YENYIPADSFIVHEDLSPELAEYLIMLDNNKKYLSYFWMRDPTVNLPRFWSHA
CLACDHVKRHOEKSKGNLEKWFNN"

CDS

BASE COUNT 515 a 336 c 320 g 534 t
ORIGIN

Query Match 85.7%; Score 783.8; DB 10; Length 1705;
Best Local Similarity 91.0%; Pred. No. 3.6e-196;

Matches 833; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1 ACCAAACTGATATTTTATGAACACTATCTGTGTGGTGTGGCATTTGGCGAG 60
DB 227 ACAAAACTGATATTTTATGAACACTATCTGTGTGGTGTGGCATTTGGCGAG 286
QY 61 ACCTTGACCTTACATCTGCGCAAGCAAGTGTCAACATCCAAAGATCGCAACAG 120
DB 287 ACCTTGACCTTACATCTGCGCAAGCAAGTGTCAACATCCAAAGATCGCAACAG 346
QY 121 GACCGTCACTGACCAACAAATCCCAAGCTGTGATCCATCCAGACATCACTTGG 180
DB 347 GACCGTCACTGACCAACAAATCCCAAGCTGTGATCCATCCAGACATCACTTGG 406
QY 181 GATCTGCAAAATTTACCTGACCAAGCTAGCGCAACCTTCGAGAAATGATTTGGATGAA 240
DB 407 GATCTGCAAAATTTACCTGACCAAGCTAGCGCGCCCTTCGAGAAATGATTTGGATGAA 466
QY 241 TTGGAATCACCACATCTCACCTCCCAAAAGAGTGGCATTTGAGCACTTGTTAACCTGACT 300
DB 467 CTAAAGTCTCCCACTCATACCCCCCAAAAGAGTGGCATTTGATCAACCTGACT 526
QY 301 CTGACCTTACCGCGTGTGATGATATTCAGATGCTCTTATGGCTCTTGTGACGGTGAAGCACA 360
DB 527 CTAAAGTATGCGCGTGTGATGATATTCAGATGCTCTTATGGCTCTTGTGACGGTGAAGCACA 586
QY 361 AATGCCCTGCTGTTTGAAGGCGCAAGCAAAAGAAATTTGGTGTGGCTGTGAGTAAAC 420
DB 587 AATGCCCTGCTGTTTGAAGGCGCAAGCAAAAGAAATTTGGTGTGGCTGTGAGTAAAC 646
QY 421 TTGGAACCTTACGATGCGCAAGTCAAGTATTTACATATGAGCTTAACAAAGAAATTTGAATC 480
DB 647 TTGGAACCTTACGATGCGCAAGTCAAGTATTTACATATGAGCTTAACAAAGAAATTTGAATC 706
QY 481 CATACCTAGCGGCAAGCAATTTGAGAAATATGTCATGATTAATAAATTTGATTCCTACCAATA 540
DB 707 CATACCTAGCGGCAAGCAATTTGAGAAATATGTCATGATTAATAAATTTGATTCCTACCAATA 766
QY 541 TCTGCTGTAAATTTTATCTTCTTGTGAATAATTTCAATCCACAGAGTATACATACGAGAA 600
DB 767 TCTGCTGTAAATTTTATCTTCTTGTGAATAATTTCAATCCACAGAGTATACATACGAGAA 826

QY 601 AAGCTATACAAATGCTTTTTCGCTGCTGCTGCTACCTGTTGTTCTGGGACCACATCTAGGCA 660
DB 827 AAGCTATACAAATGCTTTTTCGCTGCTGCTGCTACCTGTTGTTCTGGGACCACATCTAGGCA 886
QY 661 AACTATGAGAATTTATTTTCAGACAGATTCATTCATTCATGATGGAAGATTTAAGCTCTCC 720
DB 887 AACTATGAGAATTTATTTTCAGACAGATTCATTCATTCATGATGGAAGATTTAAGCTCTCC 946
QY 721 AGTAGCTAGCAAGATTTCTGAAAGAAAGTGCACAAAACAAATTAAGTTATAGTTAC 780
DB 947 AGTAGCTAGCAAGATTTCTGAAAGAAAGTGCACAAAACAAATTAAGTTATAGTTAT 1006
QY 781 TTTAAGTGAAGAAAGATTTCTGAAAGAAAGTGCACAAAACAAATTAAGTTATAGTTAT 840
DB 1007 TTTAAGTGAAGAAAGATTTCTGAAAGAAAGTGCACAAAACAAATTAAGTTATAGTTAT 1066
QY 841 TTGGCTTGCATCATCTGAAAGAAAGTGCACAAAACAAATTAAGTTATAGTTATAGGAA 900
DB 1067 CTGCTTGTGATCATCTGAAAGAAAGTGCACAAAACAAATTAAGTTATAGTTATAGGAA 1126
QY 901 TGTGTTTGGATTTA 915
DB 1127 TGTGTTTGGATTTA 1141

RESULT 8
AB035906 1707 bp DNA linear VRT 05-JAN-2002
LOCUS Gallus gallus gene for CFUT9, complete cds.
DEFINITION AB035906
ACCESSION AB035906.1 GI:18146865
VERSION CFUT9.
KEYWORDS Gallus gallus brain DNA.
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE
1 (sites)
AUTHORS Kaneo, M., Nishihara, S., Kitano, T., Narimatsu, H. and Saitou, N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL Unpublished
2 (bases 1 to 1707)
AUTHORS Kaneo, M., Saitou, N. and Kitano, T.
TITLE Direct Submission

Submitted (17-DEC-1999) Mika Kaneo, National Institute of
Genetics, Laboratory of Evolutionary Genetics, Yata 111, Mishima,
Shizuoka 411-8540, Japan (E-mail: mkaneo@med.id.yamagata-u.ac.jp,
Tel: 81-559-81-6790, Fax: 81-559-81-6789)
Sequence updated (29-Feb-2000).

COMMENT

FEATURES

source

1. 1707

/organism="Gallus gallus"

/db_xref="taxon:9031"

/tissue_type="brain"

268. 1347

/gene="cfuc-TIX"

268. 1347

/gene="cfuc-TIX"

/codon_start=1

/product="cfut9"

/protein_id="BAB82489.1"

/db_xref="GI:18146865"

/translation="MSTSKGILRPFLIVCYTLACPMACILYIKPTSNVSPWESA
SAYLKKNFESTKDYENETTLVWMPGQFDTSCQAMPNIGCHLTDRLYRK
SHAVLHHRDIDMDLNLPOQARPPQKIMANLSPHTPOKSEIHLNLTLYRK
DSDIQYPIGFLVSTNPEYEVPSKEKLVCWVSNWNEHARVKYINELSLYIRY
GOAFGEYVNDKMLIPTISCKFLYSESKELVSIKDYITTELYNALAGSPVVLGSRN
YENYIPADSFIVHEDLSPELAEYLIMLDNNKKYLSYFWMRDPTVNLPRFWSHA
CLACDHVKRHOEKSKGNLEKWFNN"

BASE COUNT 517 a 338 c 329 g 521 t 2 others
ORIGIN
Query Match 72.8%; Score 666; DB 5; Length 1707;

Best Local Similarity 83.0%; Pred. No. 4.3e-165;
Matches 759; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

OY	2	CCAAACGTGATTTTAAATGAACATCTATTCCTGGTGGGAGTGGCCATTGGGCGAGA	61
Db	434	CCAAACGTGATTTTAAATGAACATCTATTTGGTTGGCTTGGCCATTGGGCGAGA	493
OY	62	CCCTTGAACCTTACATCCCGCAAGCAATGTTCACATCCAGAGATGCCATCTCAACGG	121
Db	494	CATTCATCTTACATCCCGCAAGCAATGTTCACATCCAGAGATGCCATCTCAATGG	553
OY	122	ACCGCTACCTGATCCCAATCCAGATTCGATTCATCCAGATCCAGATCCAGATTCG	181
Db	554	ACCGCTACCTGATCCCAATCCAGATTCGATTCATCCAGATCCAGATCCAGATTCG	613
OY	182	ATCTGCAAAATTTACCTCCAGCAAGTACGACCTTCCAGAAATGGATTTGGATTAAT	241
Db	614	ATCTGCAAAATTTACCTCCAGCAAGTACGACCTTCCAGAAATGGATTTGGATTAAT	673
OY	242	TGGATTCACCACTCACTCCCAAAAGAGTGGCATTTGAGCATTTGTTTACCTGACTC	301
Db	674	TGGATTCACCACTCACTCCCAAAAGAGTGGCATTTGAGCATTTGTTTACCTGACTC	733
OY	302	TGACCTTACCGCGGTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC	361
Db	734	TGACCTTACCGCGGTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC	793
OY	362	ATCCCTTCTGTTTGAAGTCCCAAGCAAGAAATTTGGTGGTGGTGGTGGTGGTGGT	421
Db	794	ATCCCTTCTGTTTGAAGTCCCAAGCAAGAAATTTGGTGGTGGTGGTGGTGGTGGT	853
OY	422	GGAACCTTGAGCAATGCGAGATCAATGATTCATGATTCATGATTCATGATTCATG	481
Db	854	GGAACCTTGAGCAATGCGAGATCAATGATTCATGATTCATGATTCATGATTCATG	913
OY	482	ATACCTACGCGGCAAGATTTGGAGATTCATGATTCATGATTCATGATTCATGATTC	541
Db	914	ATACCTACGCGGCAAGATTTGGAGATTCATGATTCATGATTCATGATTCATGATTC	973
OY	542	CTGCTGTGAAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT	601
Db	974	CTGCTGTGAAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT	1033
OY	602	AGCTATACAAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	661
Db	1034	AGCTATACAAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1093
OY	662	ACTATGAGATTTATTTCCAGAGATTCATGATTCATGATTCATGATTCATGATTCATG	721
Db	1094	ACTATGAGATTTATTTCCAGAGATTCATGATTCATGATTCATGATTCATGATTCATG	1153
OY	722	GTGAGCTAGCAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	781
Db	1154	GAGAGCTGAGAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1213
OY	782	TTRACTGAGAGAGAGATTCGATTCATGATTCATGATTCATGATTCATGATTCATGAT	841
Db	1214	TCAATGAGAGAGAGATTCGATTCATGATTCATGATTCATGATTCATGATTCATGAT	1273
OY	842	TGCTTTCGATTCATGAG	901
Db	1274	TTCCTTTCGATTCATGAG	1333
OY	902	GGTTTGGAGATTA 915	
Db	1334	GGTTTGGAGATTA 1347	

RESULT 9
AB035905
LOCUS AB035905 1080 bp DNA linear VRT 05-JAN-2002
DEFINITION Xenopus laevis gene for xfrp9, complete cds.
ACCESSION AB035905

VERSION AB035905.1 GI:18146863
KEYWORDS xfrp9.
SOURCE Xenopus laevis DNA.
ORGANISM Xenopus laevis
REFERENCE 1 (sites)
AUTHORS Kaneko,M., Nishihara,S., Kitano,T., Nishimatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL unpublished
AUTHORS 2 (bases 1 to 1080)
TITLE Kaneko,M., Saitou,N. and Kitano,T.
JOURNAL Direct Submission
COMMENT Submitted (17-DEC-1999) Mika Kaneko, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Yata 1111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6783)
FEATURES
source
1..1080
/organism="Xenopus laevis"
/db_xref="taxon:8355"
1..1080
/gene="xfrp-TIX"
/gene="xfrp-TIX"
/protein_id="BAB82488.1"
/protein_id="BAB82488.1"
/db_xref="GI:18146864"
/translation="MWSASKGLIRAFVAVSVGLICPMICLILYVPTNMLSPLESIA
HSLVKKKNEFSSKONYNETVYLIWVPEGQFELEKSCOTLENIGCHLTDRLAYNK
SHAVLLHHDIDSMILNLPOLRPPFKIMNLSLPTPOKSEIHFENLTLYR
DSDIQVPGYGVNSTRKPRDEYPSDKPCWVSWNNEHAYKYNELIKETIITY
GQAREYLSDKLPTLASCKYLSFENSIHDIYTERKYNALNLAGSPVIVGPREN
YENYIPADSEIHEDELTRELDHILMLKIDTDEYLSYNNRKHFTVMPFWESHA
CLACDVHNRHORYSVGNLEKWFNN"
BASE COUNT 332 a 231 c 202 g 315 t
ORIGIN

Query Match 63.5%; Score 581.2; DB 5; Length 1080;
Best Local Similarity 77.2%; Pred. No. 1e-142;
Matches 706; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

OY	2	CCAAACGTGATTTTAAATGAACATCTATTCCTGGTGGGAGTGGCCATTGGGCGAGA	61
Db	167	CTTAAACAGATTTTAAATGAACATCTATTCCTGGTGGGAGTGGCCATTGGGCGAGA	226
OY	62	CCCTTGAACCTTACATCCCGCAAGCAATGTTCACATCCAGAGATGCCATCTCAACGG	121
Db	227	CTTGAACCTTACATCCCGCAAGCAATGTTCACATCCAGAGATGCCATCTCAACGG	286
OY	122	ACCGTTCACCTGACCAACCAATCCAGATTCGATTCATCCAGAGATCCAGATTCGAT	181
Db	287	ACCGTTCACCTGACCAACCAATCCAGATTCGATTCATCCAGAGATCCAGATTCGAT	346
OY	182	ATCTGCAAAATTTACCTCCAGCAAGTACGACCTTCCAGAAATGGATTTGGATTAAT	241
Db	347	ATCTGCAAAATTTACCTCCAGCAAGTACGACCTTCCAGAAATGGATTTGGATTAAT	406
OY	242	TGGAATTCACCACTCACTCCCAAAAGAGTGGCATTTGAGCATTTGTTTAACTGACTC	301
Db	407	TGGAATTCACCACTCACTCCCAAAAGAGTGGCATTTGAGCATTTGTTTAACTGACTC	466
OY	302	TGACTTACCGCGGTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC	361
Db	467	TGACTTACCGCGGTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC	526
OY	362	ATCCCTTCTGTTTGAAGTCCCAAGCAAGAAATTTGGTGGTGGTGGTGGTGGTGGT	421
Db	527	ATCCCTTCTGTTTGAAGTCCCAAGCAAGAAATTTGGTGGTGGTGGTGGTGGTGGT	586

Db 712 GAGACTCAGACGACAAAGACTATACACTGAGAGAGCTCGGAGGATCGCTGTCCT 771
Oy 625 GGCTCTGATCTGTTGTTGTTGGACCATCTAGGAAACATATGATATATATCCAGCA 684
Db 772 GGGACATCCAGATTTTCTGGGGGACACCGAGAAACATATAGACATCTTGCCCTT 831
Oy 685 GATTCTATTCATTCATGGAAGATTTAACTCTCCAGTGAGTACGAAATGATCTGAG 744
Db 832 GACTCTTCATCCACGTCGATGACTTGGCCAGTGTGTAACGTGCGCAGTACTGAC 891
Oy 745 GAGTGCAGAAAACATATAGTATACCTTATACCTTATACGAGGAG 795
Db 892 GAACTAGAGGAGACACCGAGAAAGTACCAAGCTACTCTCCAGTGGCAGG 942

RESULT 15
GGU73678 2657 bp DNA linear VRT 20-DEC-1996
LOCUS DEFINITION Gallus gallus alpha-1,3-fucosyltransferase (CFT1) gene, complete
ACCESSION U73678
VERSION U73678.1 GI:1657998
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2657)
AUTHORS Lee, K.P., Carlson, L.M., Woodcock, J.B., Ramachandra, N., Schult, T.L., Davis, T.A., Lowe, J.B., Thompson, C.B. and Larsen, R.D.
TITLE Molecular cloning and characterization of CFT1, a developmentally regulated avian alpha(1,3)-fucosyltransferase gene
JOURNAL J. Biol. Chem. 271 (51), 32960-32967 (1996)
MEDLINE 97115837
REFERENCE 2 (bases 1 to 2657)
AUTHORS Lee, K.P., Carlson, L.M. and Larsen, R.D.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1996) Immune Cell Biology Program, Naval Medical Research Institute, 8901 Wisconsin Ave., Bethesda, MD 20889, USA
FEATURES
source 1..2657
gene /organism="Gallus gallus"
/db_xref="taxon:9031"
170..1240
/gene="CFT1"
170..1240
/gene="CFT1"
/note="similar to mammalian fucosyltransferase FucTIV"
/codon_start=1
/product="alpha-1,3-fucosyltransferase"
/protein_id="AAC60060.1"
/db_xref="GI:1657998"
/translation="MEIGPRMSPAPRPPGPRRRRRALLGLALALALYLCVRELR
RHSAGAREGEVTLVLMWEPGRPWPRADCKRRYNITGCLLSAQRGRGAEAVLFH
HNDLALHGGDLPGRPPRQWRVWMAFESPSPGIRGLAGLFNMTYSRSDV
FVYGYLVEPSPRPFVLRKSLVAVYVSNNEBHARVYRQLEHPIVYVGARG
MALGSAVYKAYKPYLAENSOHTDITKLMKNAPAAVAPVYLGPRANERE
IPADSPFIHVDPPRLATYIKFLDKNKPSTRRYFAMNKEVHTSWDEHYCKVC
EAVRTAGNOLKTVQNDAGWFS"
BASE COUNT 581 a 712 c 818 g 546 t
ORIGIN

Query Match 19.3%; Score 176.2; DB 5; Length 2657;
Best Local Similarity 53.1%; Pred. No. 7.3e-36;
Matches 451; Conservative 0; Mismatches 383; Indels 15; Gaps 3;

Oy 28 ACTATCTCTGTTGGTGGCCATTGGGACAGCCTTGACCTTACATCCGCAAGCA 87
Db 341 ACGGTGCTGTGTGGTGGGACCTTGCGGCCCTGGGCGCCGACGACGCGAGG 400
Oy 88 ATGTTCAACTCCAGAGGACATCTCAGACGCGCTTCACTGACCAAAATCCCAT 147

Db 401 CGCTAACATCAAGGAGGCTGCTCTGAGCGCCGACCGGGGGGAGCGGAGGCTCG 460
Oy 148 GCAGTCTGATCCATCAGACGAGACATCAGTGGATCTGCACAAATTTACTCGAAG-- 205
Db 461 GCGGTGCTCTCCACACAGCGGACATCGGCGCTGCAGCGCCGACAGGGGCTGCCCGGG 520
Oy 206 -----CTAGGCCACCCCTTCCAGAAATGATTTGGATTTGGAATCCAACTCAC 258
Db 521 CCCCCCGCGACCCCCGCGGACAGCGGTGGTGTGATGAATTCGATCGCCCTCGCAC 580
Oy 259 ACTCCCAAAAAGATGAGCATGTAGCACTTTTAACTGACTGTACTACCGCGGTGAT 318
Db 581 TCCCGCGGCTGCGGGGCGCTGCGGGGCTTCTTAACTGACCAATGTCACACGAGAC 640
Oy 319 TCAGATATCCAGTGCCTTAGCTTTCAGCGGTAAGCAAAATCCCTTGATTTGAA 378
Db 641 TCGGAGCTGTCTGCTGCTTACGGGTACCTTACGAGCGCGCGCGGCGGCTTCTGTG 700
Oy 379 GTGCCAAGCAAGAGAAATTTGTGCTGGTTGTGAGTAATGAAACCTTGACATGCC 438
Db 701 CTGCTGCGCAAAAGCGGCGTGGTGGCTGGGTATCAGCAATGGAAGAGAGCAGCGC 760
Oy 439 AGAGTCAATTTACATAGCTAAGCAAGCAAAAGCATTGAATCCATACCTACGGCAAGCA 498
Db 761 CGGCTGCGCTACTACCGGCGAGCTGAAGAGCACCTGCCCATGACGTTGACGGG--GCG 817
Oy 499 TTGAGAAATATGTCATATGAATAAATTTGATTCCTACCATATCTGCTGTAAATTTAT 558
Db 818 CGGGGATGGCGCTGCTGGAGGCGAGCTGTGTAAAGCGGTGCGCTTCAAGTTCTAC 877
Oy 559 CTTTCTTTGAAAATTCATATCACAAGATTTACATCAGGAAAGCTAT--ACAATGCT 615
Db 878 CTGCGCTTGAAGAATCCGACAGACGAGACTACATCAGGAGAGGTGTGAAGAACGCC 937
Oy 616 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Db 938 TTGCGCGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Oy 676 ATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 735
Db 998 ATCCCGCGGACCTCTTCACTCCAGCTGAGCATTTCCCGGACCGGCGGCTGCGGAC 1057
Oy 736 TATCTGAGGAGAGTGCAGCAAAAACAATTAAGTTAACTTAACTTAACTGAGGAAG 795
Db 1058 TACCTGAAGTTCCTCGATAAAACAAGCCAGCTACAGAGGATATTTCGCTGCGGAA 1117
Oy 796 GATTCACTGAATCTTCCAGCATTTTGGGAATTCACATGATGTTGGCTTGCATCAT 855
Db 1118 AAGTATGAAGTCCACGTCACGCTTCTTCTGGATGAGCATTAATGCAAGGTTTCCGAGCC 1177
Oy 856 GTGAAAAG 864
Db 1178 GTGAGAGCG 1186

Search completed: October 6, 2002, 04:33:24
Job time : 1341.59 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 102.682 Seconds
(without alignments)
15299.444 Million cell updates/sec

Title: US-09-744-748-4_COPY_1454_2368

Perfect score: 915
Sequence: 1 accaaactgattatttaa.....agaatggttggaattaa 915

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : N_Geneseq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	915	100.0	2676	21	AAZ92647 Human alpha-1,3-fu
2	915	100.0	2822	21	AAZ92646 Human alpha-1,3-fu
3	791.8	86.5	2036	21	AAZ92672 Murine alpha-1,3-f
4	791.8	86.5	2170	21	AAZ92645 Murine alpha-1,3-f
5	127.4	13.9	1814	18	AAZ59505 Murine myeloid-11n
6	127	13.9	3594	18	AAZ85219 Mouse alpha-fucosy
7	114.4	12.5	1701	15	AAO77732 Human alpha-1,3-fu
8	114.4	12.5	1701	21	AAZ21136 Human low adenosin
9	114.4	12.5	1701	21	AAZ5014 Human adenosine re

10	114.4	12.5	6944	21	AAZ21137
11	114.4	12.5	6944	21	AAZ35015
12	113.4	12.4	1654	21	AAO56905
13	113.4	12.4	1654	18	AAZ61680
14	112.2	12.3	1086	15	AAO56912
15	112.2	12.3	2042	18	AAZ80111
16	112.2	12.3	2043	12	AAO13330
17	112.2	12.3	2043	15	AAO56906
18	112.2	12.3	2043	18	AAZ76769
19	112.2	12.3	2043	18	AAZ61675
20	111.8	12.2	1126	21	AAZ21135
21	111.8	12.2	1126	21	AAZ35013
22	111.8	12.2	1316	15	AAO56911
23	111.8	12.2	1316	18	AAZ61679
24	103.6	11.3	795	17	AAZ13798
25	102.4	11.2	2175	11	AAO06691
26	100.8	11.0	1256	21	AAZ21134
27	100.8	11.0	1256	21	AAZ35012
28	100.8	11.0	1400	13	AAO31436
29	100.8	11.0	1488	15	AAO56910
30	100.8	11.0	2175	12	AAO14382
31	100.8	11.0	2175	24	AAZ17082
32	100.8	11.0	2861	12	AAO14383
33	100.8	11.0	2861	21	AAZ21133
34	100.8	11.0	2861	21	AAZ35011
35	100.8	11.0	2861	24	AAZ17083
36	100.8	11.0	3647	12	AAO13333
37	100.8	11.0	3647	15	AAO56909
38	100.8	11.0	3647	15	AAZ61678
39	97.6	10.7	2134	18	AAZ59506
40	66.6	7.3	1578	23	ABL11867
41	66.6	7.3	4229	23	ABL11866
42	52.2	5.7	1353	21	AAZ51685
43	44.4	4.9	676	22	AAZ41344
44	44.4	4.9	1479	24	AAO5334
45	44.4	4.9	2557	24	AAO5333

ALIGNMENTS

RESULT 1	AAZ92647	standard; cDNA; 2676 BP.
ID	AAZ92647	
XX	AAZ92647	
AC	AAZ92647	
DT	05-JUN-2000	(first entry)
XX		
DE	Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.	
XX		
KW	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;	
KW	brain; kidney; recombinant expression; transgenic animal; knockout	
KW	animal; FDC-TIV; drug screening; inhibitor; potentiators; diagnosis;	
KW	treatment; cancer; human; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	295..1374
FT		/tag= a
FT		/product= "Human alpha-1,3-fucosyltransferase"
PN	MO200006708-A1.	
XX		
PD	10-FEB-2000.	
XX		
PF	29-JUL-1999;	99WO-JP04092.
XX		
PR	29-JUL-1998;	98JP-0213823.
XX		
PA	(KYO) KYOWA HAKKO KOGYO KK.	
XX		

Human low adenosin
Human adenosine re
pcDNA1-alpha-(1-3)
Human alpha(1,3)-f
Lewis blood group
Fucosyltransferase
Glycosyltransferas
DNA encoding a gly
Human alpha(1,3)/
Human alpha(1,3)/
Human low adenosin
Human adenosine re
DNA encoding a gly
Human alpha(1,3)-f
Alpha-(1,3/1/4)-fu
CDX, a MILA involv
Human low adenosin
Human adenosine re
Encodes a HeLa cel
DNA encoding a gly
Clone 7.2 encoding
Human cDNA clone 7
Clone 1 encoding 1
Human low adenosin
Human adenosine re
Human cDNA clone 1
GDP-Fuc:beta-D-Gal
DNA encoding a gly
Human alpha(1,3)-f
Human myeloid deri
Drosophila melanog
C. elegans alpha-1
cDNA encoding nove
Human fucosyltrans
Human fucosyltrans

CC Invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention vectors additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA292645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA80995), and sequences AA292670-292673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).

XX Sequence 2036 BP: 666 A; 365 C; 348 G; 657 T; 0 other:

Query Match 86.5%; Score 791.8; DB 21; Length 2036;
 Best Local Similarity 91.6%; Pred. No. 3.6e-219;

Matches 838; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 ACCAAACGATGATTTTATATGAACTCTTCTGGTGGGTGGCCATTGGGCGAG 60
 DB 174 ACAAAACGATGATTTTATATGAACTCTTCTGGTGGGTGGCCATTGGGCGAG 233
 QY 61 ACCTTTGACCTTACATCTCCGCAAGCAATGTTCAACATCCAGAGTCCATCTTCACAAAG 120
 DB 234 ACCTTTGACCTTACATCTCCGCAAGCAATGTTCAACATCCAGAGTCCATCTTCACAAAG 293
 QY 121 GACCGTTCACTGATACAAACAAATCCATGCTGATGATCAGCCAGCATCACTGTTGG 180
 DB 294 GACCGCTCACTGATACAAACAAATCCATGCTGATGATCAGCCAGCATCACTGTTGG 353
 QY 181 GATCTGACAAATTTTACCTAGCAAGCTAGGCGACCTTCGCAAGAAATGGATTGGATGAT 240
 DB 354 GATCTGACAAATTTTACCTAGCAAGCTAGGCGACCTTCGCAAGAAATGGATTGGATGAT 413
 QY 241 TTGGATGACCAAGTACACCTCCCAAGAGTGGATGAGCACTGTTTAACTGAGT 300
 DB 414 TTGGATGACCAAGTACACCTCCCAAGAGTGGATGAGCACTGTTTAACTGAGT 473
 QY 301 CTGACTTACCGCGGATTCAGATATCCAGTGGCTTATGGCTTCTTGACGGTACGACA 360
 DB 474 CTGACTTATGCGCGGATTCAGATATCCAGTGGCTTATGGCTTCTTGACGGTACGACA 533
 QY 361 AATCCCTTGCTGTGAAAGTGCAGCAAGAGAAATGGTGTGCTGGATTGTGAGTAAC 420
 DB 534 AATCCCTTGCTGTGAAAGTGCAGCAAGAGAAATGGTGTGCTGGATTGTGAGTAAC 593
 QY 421 TGGAACTTGCATGACAGATGCAAGTATTACATGAGTAAGCAAGCAATTTGAAATC 480
 DB 594 TGGAACTTGCATGACAGATGCAAGTATTACATGAGTAAGCAAGCAATTTGAAATC 653
 QY 481 CATACCTAGGGGCAAGCATTTGGAGATATGTCAATGATTAATAATTTGATCTCCATTA 540
 DB 654 CATACCTAGGGGCAAGCATTTGGAGATATGTCAATGATTAATAATTTGATCTCCATTA 713
 QY 541 TCGGCTGTAATTTTATCTTCCCTTGAATTCATCCACAAGATTTACATCAGGAA 600
 DB 714 TCGGCTGTAATTTTATCTTCCCTTGAATTCATCCACAAGATTTACATCAGGAA 773
 QY 601 AAGCTATACATGCTTTCTGGCTGGCTCTGTACCTGTTCTTGCGGACCATCTAGGAA 660
 DB 774 AAGCTATACATGCTTTTGGCTGGCTCTGTACCTGTTCTTGCGGACCATCTAGGAA 833

QY 661 AACTATGAGAAATATATTCACAGATTCATTCACTTCATGAGAAATATTAAGTCTGCC 720
 DB 834 AACTATGAGAAATATATTCACAGATTCATTCACTTCATGAGAAATATTAAGTCTGCC 893
 QY 721 AGGAGCTGCAAGATGATCTGAAGAGTCCGCAAGAAACAAATAGTATACCTTACTTAC 780
 DB 894 AGGAGCTGCAAGATGATCTGAAGAGTCCGCAAGAAACAAATAGTATACCTTACTTAC 953
 QY 781 TTTAAGTGAAGAGATTTCACTGTAATCTTCACAGATTTGGGATCATCATGATGCT 840
 DB 954 TTTAAGTGAAGAGATTTCACTGTAATCTTCACAGATTTGGGATCATCATGATGCT 1013
 QY 841 TTGGCTTGCATCATGTAAGAAAGCATCAAGAAATATAAGTCTGTTGTAATTTAGAGAA 900
 DB 1014 TTGGCTTGCATCATGTAAGAAAGCATCAAGAAATATAAGTCTGTTGTAATTTAGAGAA 1073
 QY 901 TGGTTTGGAAATTA 915
 DB 1074 TGGTTTGGAAATTA 1088

RESULT 4

AA292645
 ID AA292645 standard; cDNA; 2170 BP.

AC AA292645;

XX 05-JUN-2000 (first entry)

DE Murine alpha-1,3-fucosyltransferase cDNA.

XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
 KW treatment; cancer; murine; mouse; ss.

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 115..1194

FT /tag=a

FT /product="Murine alpha-1,3-fucosyltransferase"

PN MO200006708-A1.

XX 10-FEB-2000.

XX 29-JUL-1999; 99WO-TP04092.

XX 29-JUL-1998; 98JP-0213823.

PA (RYOW) KYOMA HAKKO KOGYO KK.

PI Narimatsu H, Kudo T, Sasaki K;

XX WPI: 2000-183120/16.

DR P-PDB: AAY80995.

PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x

PT epitope and an antibody recognizing it useful for diagnosis of brain

PS and kidney disease and cancer.

XX Claim 4; Page 127-134; 172pp; Japanese.

CC The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylgalactosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising the
 CC these DNA sequences. The invention additionally encompasses the

CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA292645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA90995), and sequences AA292670-292673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).

XX Sequence 2170 BP: 715 A; 395 C; 368 G; 692 T; 0 other;

Query Match 86.5%; Score 791.8; DB 21; Length 2170;

Best Local Similarity 91.6%; Pred. No. 3.7e-219;

Matches 838; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

OY 1 ACCAAAGTATTTATTTATGAACTACTATCTGTTGGTGGTGGCCATTGGGCGAG 60
 DB 280 ACAAAGTATTTATTTATGAACTACTATCTGTTGGTGGTGGCCATTGGGCGAG 339
 OY 61 ACCTTGACCTTACATCTGCGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACACG 120
 DB 340 ACCTTGACCTTACATCTGCGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACACG 399
 OY 121 GACCGTCACTATACAAATCCATCGATTCGATTCATCACCAGAACATCACTATGG 180
 DB 400 GACCGTCACTATACAAATCCATCGATTCGATTCATCACCAGAACATCACTATGG 459
 OY 181 GATCTGACAAATTTTACCTCAGCAAGTAGGCAACCTTCAGAAATGATTTGGATGAT 240
 DB 460 GATCTGACAAATTTTACCTCAGCAAGTAGGCAACCTTCAGAAATGATTTGGATGAT 519
 OY 241 TTGGAATCAACCACTACACTCCCAAAAAGAGTGGCATTTGAGCACTGTTTAACTGACT 300
 DB 520 TTGGAATCAACCACTACACTCCCAAAAAGAGTGGCATTTGAGCACTGTTTAACTGACT 579
 OY 301 CTGACTTACCGCGCGATTCAGATATCCAAAGTCCCTATGCGCTCTTGAACGGTAACACA 360
 DB 580 CTGACTTACCGCGCGATTCAGATATCCAAAGTCCCTATGCGCTCTTGAACGGTAACACA 639
 OY 361 AATCCCTTGGTGTGAAGTGGCAAGCAAGAAATTTGGTGTGGTGTGAGTATAC 420
 DB 640 AATCCCTTGGTGTGAAGTGGCAAGCAAGAAATTTGGTGTGGTGTGAGTATAC 699
 OY 421 TGGAAACCTGAGCAAGCAAGTCAAGTATTAACAATGACCTTAACCAAAACATTTGAATC 480
 DB 700 TGGAAACCTGAGCAAGCAAGTCAAGTATTAACAATGACCTTAACCAAAACATTTGAATC 759
 OY 481 CATTAAGGCAAGCAAGTATTAACAATGATTAACAATTTGATTTCTTACCATTA 540
 DB 760 CATTAAGGCAAGCAAGTATTAACAATGATTAACAATTTGATTTCTTACCATTA 819
 OY 541 TGTGCTTAATTTATTTCTTCTTGAATTTCAATCCACAAGATTTACATCAGCA 600
 DB 820 TGTGCTTAATTTATTTCTTCTTGAATTTCAATCCACAAGATTTACATCAGCA 879
 OY 601 AAGCTATACAAATGCTTTTGGCGGCTGTAAGTCTGTTCTTGGAGCAATCTAGGAA 660
 DB 880 AAGCTATACAAATGCTTTTGGCGGCTGTAAGTCTGTTCTTGGAGCAATCTAGGAA 939
 OY 661 AACTATGAAATTTATTTCCAGCAAGTATTAACAATTTTAACTCTCCG 720
 DB 1180 AACTATGAAATTTATTTCCAGCAAGTATTAACAATTTTAACTCTCCG 720

DB 940 AACTATGAAATTTATTTCCAGCAAGTATTAACAATTTTAACTCTCCG 999
 OY 721 AGTACGCTACCAAGTATTTGAAAGTACGACAAACAAATATGATTAACCTAGTAC 780
 DB 1000 AGTACGCTACCAAGTATTTGAAAGTACGACAAACAAATATGATTAACCTAGTAC 1059
 OY 781 TTTAACTGGAAGAAAGATTTTCACTGTAATCTTCACAGATTTTGGAAATCAGATGAT 840
 DB 1060 TTTAACTGGAAGAAAGATTTTCACTGTAATCTTCACAGATTTTGGAAATCAGATGAT 1119
 OY 841 TTGGCTTGGGATCATGTTGAAAGGCAATCAAGAAATATTAAGTCTGTTATTAAGGAA 900
 DB 1120 TTGGCTTGGGATCATGTTGAAAGGCAATCAAGAAATATTAAGTCTGTTATTAAGGAA 1179
 OY 901 TGGTTTGGAAATTA 915
 DB 1180 TGGTTTGGAAATTA 1194

RESULT 5

AAT59505
 ID AAT59505 standard; cDNA; 1814 BP.

XX AAT59505;

XX 06-MAY-1997 (first entry)

XX Murine myeloid-1 lineage alpha-(1,3)-fucosyltransferase cDNA.

XX Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM;
 XX septic shock; septicemia; therapy; ss.

XX Mus sp.

XX Key Location/Qualifiers

XX CDS 325..1353

XX W09640881-A1.

XX 19-DEC-1996.

XX 08-MAY-1996; 96MO-US06427.

XX 07-JUN-1995; 95US-0483151.

XX (GENE) GEN HOSPITAL CORP.

XX Holgersson J, Seed B;

XX WPI; 1997-108639/10.

XX P-PSDB; AAW11820.

XX New murine alpha-(1,3)-fucosyltransferase - for fucosylating an
 XX antibody to protect mammals against e.g. septic shock or septicemia

XX Claim 6: Fig 3: 58pp; English.

XX A cDNA clone (AAT59505) codes for murine myeloid lineage alpha-
 XX (1,3)-fucosyltransferase (AAW11820), an enzyme which has a relatively
 XX strict substrate requirement for sialylated N-acetylglucosamine,
 XX which can account for the presence of the sialyl-Le-x epitope on
 XX murine cells, and which is more effective than Fuc-TIV in support
 XX of E-selectin-mediated COS cell adhesion. It was isolated from a
 XX murine myeloid 32D cl3 cell cDNA library by identification of a
 XX clone capable of directing the expression of sialyl-Le-x
 XX determinants. Transformed host cells (e.g. 32D cl3 or human 293
 XX cells) expressing the alpha-(1,3)-fucosyltransferase, and pref.
 XX another fucosyltransferase such as human Fuc-TIV (see also AAT59506),
 XX can be used to fucosylate an antibody (e.g. IgG or IgM) for use in
 XX protecting an animal against an adverse immune reaction, esp. septic
 XX shock or septicemia.

```

SO Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other;
Query Match 13.9%; Score 127.4; DB 18; Length 1814;
Best Local Similarity 52.1%; Pred. No. 1,3e-26;
Matches 334; Conservative 0; Mismatches 301; Indels 6; Gaps 2;

OY 104 GATGCCATTCACACAGGACCGCTTCACTGACACAAATCCATGCATGCTGATCCATC 163
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 GCTGCCGCTGAGTGCATACCGGACGCTGACGACATGCTGATGCTGCTGCTCCACC 607
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 164 ACCGAGACATGTTGGAGATCTGACAAATTTACCTGACGACGACGACGACGACGAC 223
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 ACCGAGACATGTTGGAGATCTGACAAATTTACCTGACGACGACGACGACGACGAC 667
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 224 AATGATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 CTGGGCTGCGGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 727
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 284 ACTGTTAACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 343
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 GATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 344 TCTTGACGCTGACACAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 403
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 GCTTGAGACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 404 GCTGGGTTGTAGTAACTGACACCTGACGATGACGATGACGATGATGATGATGATG 463
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CCGTGGTGTATGCAATTTCCAGAGGACGACGACGACGACGACGACGACGACGACG 904
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 464 GCAAAAGCATTTGAATCATCTACCTACGACGACGACGACGACGACGACGACGACG 523
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 CCCCCTCATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 524 ATTGATTCCTACCATATCTCTGTAATTTATCTCTGTAATTTATCTCTGTAATTT 583
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 GTCTCTCTCCACATTTGGCCCGGACGCTTCACTGACGCTTGAACATCAGACATC 1024
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 584 AGGATTCATACGAGGAAAGTAT--ACAATGCTTTTCTGCTGCTGCTGCTGCTGCT 640
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 GGGACTTACATCTGAGAAAGTCTGCGCATGCTGCGGCTGGGCTGGGCTGACCCG 1084
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 641 TTCTGGGACATCTAGGAAACTATGAGAAATTAATTAATTAATTAATTAATTAAT 700
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 CGCTGGGACCTCTGCGGACGCTGACGAGGCTTTTGGCCACGACATGCTTTGTAC 1144
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 701 TGAAGATTTATTAATCTCCAGAGCTAGCAAAAGTATCTG 741
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1145 TGAAGATTTATTAATCTCCAGAGCTAGCAAAAGTATCTG 1185
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
AAT85219 standard; DNA: 3594 BP.
ID AAT85219
XX AAT85219;
AC
XX
XX 02-MAR-1998 (first entry)
DE
XX
XX Mouse alpha-fucosyltransferase Fuc-TVIII gene.
XX
XX Alpha-fucosyltransferase; Fuc-TVIII gene; mouse;
XX stably Lewis x determinant; oligosaccharide; antiinflammatory;
XX inhibitor; ds.
XX
XX Mus musculus.
XX
XX
XX Key Location/Qualifiers
XX CDS 996..3082
XX FT /tag- a
XX FT /note= "includes introns"
XX FT exon 465..477
XX FT /tag- b

OY 104 GATGCCATTCACACAGGACCGCTTCACTGACACAAATCCATGCATGCTGATCCATC 163
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2277 GCTGCCGCTGAGTGCATACCGGACGCTGACGACATGCTGATGCTGCTGCTCCACC 2336
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

OY	164	ACCAGACATCAGTGGGATCGTCAACAATTTCCTCAGACACTGGCCACCTTCACGA	223
Db	2337	ACCGTGAGCTGCAAACCCGGCAATCTCTCTACCTCCGTGACCAAGGCCACAGCACGC	2386
OY	224	AATGATTTGGATGAATTGGGAATCAACAACACTCACACTCCCACAAAAGATGGCATTTAGC	283
Db	2397	CTTGGGCTCGGGCCCTCCATCGAATTCGCCACAGTAATACCATGTCGTCCATCGCTTCGGG	2456
OY	284	ACTTGTTTAACCTGCAGCTCTGACTTACCGCCGGATTCAGATATCCAAAGTGCCTTAGGCT	343
Db	2457	GCACTCTTCAACTGGGCTGAGGTATACGGCGATGATTCAGATATCTTTGTACCTACGGTTC	2516
OY	344	TCTTGACGTACACACAATACCTTGTGTGTGAATGCGCAAGAAGAAATTGGTGT	403
Db	2517	GCTTGGAGCCTCTCTCTGGGCG---CACATCCCCACTACCGGCCCAAAAGCAGATGGCTG	2573
OY	404	GCTGGGTTGTGATTAACCTGGAACCCCTGAGCATGCCAGTCAAGTATTACATGAGCTAA	463
Db	2574	CCTGGGTGATCAGCAATTTCCAGAGACGGCAGAGAGGTGCACAAAGCTTAACGGGACGTGG	2638
OY	464	GCAAACGATTTGAATTCATACCTACGGGCMAGCATTTGGAGAAATATGTCAATGATAAA	523
Db	2634	CCCCCATCTGCAGGTGAGTGTGTGGTGGCGCCAGCGGAGGCCCCATATGCGCTAATT	2693
OY	524	ATTGATTTCCACCAATPCTGCTTGTGAATTTATCTTCTTGAATAATTCATTCACA	583
Db	2654	GTCGTGCTGCCACTTTGGCCCCGTACCGCTTCTACCTGGCCTTTAGANCTACAGCATC	2753
OY	584	AGGATTCATCAGCGAAACGAT--ACAAATGCTTTTCTGGCTGGCTGTACCTGTG	640
Db	2754	GGGACTCATCAGTGAAGATTTCTGGCGGAATGCCCTGGCGGCTGTGATGCCGTGG	2813
OY	641	TTCTGGGACCATAGGGAACATATAGAAATTAATTTACAGCAATTCATTCATTCATG	700
Db	2814	CGGTGGACCTCTCGGGCCACTACGAGGCTTTGTGTGCCACCAAGATGCTTTGTACAG	2873
OY	701	TGGAAGATTAACTCTCCAGTACGACTAGC	731
Db	2874	TGAGCACTTCAGCTCTGCCCCGTGAAGTGGC	2904
<hr/>			
RESULT	7		
ID	AAO77732		
XX	AAO77732 standard; cDNA to mRNA; 1701 BP.		
AC	AAO77732;		
XX	22-JUN-1995 (first entry)		
DE	Human alpha-1,3-fucosyltransferase cDNA.		
KW	Alpha-1,3-fucosyltransferase; sialylated Lewis blood group antigen;		
OS	Lewis x; Lewis a; glycosylation; glycoprotein; glycolipid; ds.		
FT	Homo sapiens (THP-1 cells).		
FT	Key	Location/Qualifiers	
FT	CDS	123..1151	
XX		/tag a	
XX		/product= alpha-1,3-fucosyltransferase	
XX	MO9423021-A.		
PD	13-OCT-1994.		
XX	28-MAR-1994;	94MO-JP00496.	
XX	29-MAR-1993;	93JP-0069016.	
PR	(KYOW) KYOWA HAKKO KOGYO KK.		
PA			
XX			
PI	Hanai N, Kurata K, Nishi T, Sasaki K;		

XX	WPI; 1994-333183/41.
DR	P-PsDB; AAR63215.
XX	
PT	Alpha-1,3-Fucosyl-transferase DNA and protein - useful for prodn.
PT	of active sugar chains, e.g. sialylated Lewis X
PS	Claim 3; Page 99-102; 155pp; Japanese.
XX	
CC	The sequence coding for alpha-1,3-fucosyltransferase was isolated
CC	from a human THP-1 cell cDNA library. The enzyme catalyses the
CC	addition of fucose via a 1,3-linkage to N-acetylglucosamine
CC	contained in glycoproteins and glycolipids, e.g. for sialylation of
CC	Lewis blood group antigens.
XX	
SQ	Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;
Query Match	12.5%; Score 114.4; DB 15; Length 1701;
Best Local Similarity	50.8%; Pred. No. 7.6e-23;
Matches 327; Conservative	0; Mismatches 311; Indels 6; Gaps 2.
OY	91 TTCACATCCAAAGATGCCATTCTCAACAAGACCCTTACTGTACACAAATAATCCCATGCA 150
Dd	333 TAGGGCATCGCGCGTGCCACCTGAGTGGCCAACCGAAGCCTGTGGCCAGCGCCGACGCG 392
OY	151 GTTCGTGATCCATCACCCAGACATCAATGATTGGATCTGCACAAATTTCCTCAGCAAGCTTA 210
Dd	393 GTGGCTCTCCACACCGGAGCTGCAGAACCCGGCGGTCCACCTGCCCTCGCGCCGACGCG 452
OY	211 CCACCCCTCCAAANATGATTTGGATGAATTTGSAATTCACCAATCAGACTCCCCAAAAG 270
Dd	453 CGCGCAGGCGACCCCTGGGTGTGGCTTCCTCAATGAGATCTCTTAGCCACCCACGAGCTTC 512
OY	271 ACTGGCATTTGACACTTGTGTAACTGATCGATCTGACTTACCGCGGTATTCAGATATCCAA 330
Dd	513 ACCCACTCCAGGAGCATCTTCAACTGGGGTGTGAGCTGACGCTACCGGCGGACCTGGACATCTTT 572
OY	331 GTGCCCTTATGCTTCTTGACGGTAAAGCACAAATCCCTTCGTGTTTGAAGTSCCAAGCAA 390
Dd	573 GTGCCCTTATGCTGCTGAGAGCCCACTGGGGGCGCT--CGCCACCGCTGCCACGCAAG 629
OY	391 GAGAAATGATGCTGCTGGGTGTGGAGTAACTGGAAACCCGAGACATGCCAGATCAAGTAT 450
Dd	630 ACCAGAGGTGCGCCCTGGGTGTGACCACTTCCAGAGAAGCGGACCTGCTGCTCAGGCTG 689
OY	451 TACAATGAGCTAAGCAAAAAAGCATTTGAAAATCCATTACCTACGGGCAAGCATTTGGAGAAAT 510
Dd	690 TACCGGCACTGGCCCTCATCTGGGGGTGAGATGTCTTTGGCCGTGCCAATGGAGGCCA 749
OY	511 GTCAATGATATAAAATTTGATTCCTACCAATATGCTGTGTAAATTTTATCTTTCCTTTGAA 570
Dd	750 CTGTGCGCCAGCTGCTGCTGATGCCACACCGTGGCCCAAGTACCGCTTCTACTGTCTTTGAG 809
OY	571 AATTCAATCCACAAGATATTACATCAAGGAAAAAGCTAT---ACAATGCTTTTGTGGCTGCG 627
Dd	810 AACTCTCAGACACCGGACACTCAATTCTGGGAAATTTCTGGCGACAGCACTGGTGGCTGGCG 869.
OY	628 TCTGTACCTGTTGTTCTGTGGACCACTAGGAAAACTATAGAAATATATTCCAGAGAT 687
Dd	870 ACTGTGCCAGTGTCTGTGGGCGCCCGACGGGCACTATGAGAGCCCTTCGTGCGGCTGAC 929
OY	688 TCATTCATTCATGTGGAAGATTATTAACCTCTCCAGTAGCTAGC 731
Dd	930 GCCTTCGTGCATGTGATGACTTTGGCTCAGCGCCAGAGACTGGC 973
RESULT 8	
AAF21136	
ID	AAF21136 standard; DNA; 1701 BP.
XX	
NC	AAF21136;
XX	
DT	14-MAR-2001 (first entry)

DE Human low adenosine antisense oligonucleotide related sequence #2703.
XX
XX Low adenosine antisense oligonucleotide: phosphorothioate; allergy;
XX human; airway disorder; bronchoconstriction; lung inflammation;
XX surfactant depletion; respiratory; bronchodilator; antiinflammatory;
XX immunosuppressive; antasthmatic; analgesic; hypotensive; cytostatic;
XX respiratory obstruction; pulmonary obstruction; impeded respiration;
XX surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
XX respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis
XX pulmonary hypertension; emphysema; chronic obstructive pulmonary disease (COPD);
XX chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
XX cancer; ss.
XX
XX Homo sapiens.
OS
XX WO200062736-A2.
PN
XX 26-OCT-2000.
PD
XX 24-MAR-2000; 2000WO-US08020.
PF
XX 06-APR-1999; 99US-0127958.
PR
XX (UYEC-) UNIV EAST CAROLINA.
PA (NYCE/) NYCE J W.
XX
XX NYCE JW:
PI
XX WPI; 2000-679539/66.
DR
XX
XX Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
XX Disclosure; Page 1034; 1592pp; English.
XX
XX The present invention describes low adenosine (A) content antisense
XX oligonucleotides and compositions (I) comprising them. In the antisense
XX oligonucleotides the A is replaced by a 'Universal' or alternative base
XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
XX immunosuppressive, antisthmatic, hypotensive and cytostatic activities
XX The antisense oligonucleotides and (I) can be used to down-regulate the
XX expression and or activity of target polypeptides associated with
XX lung/respiratory disorders and malignancies, such as stimulating and
XX activating peptide factors and transmitters, transcription factors,
XX immunoglobulins and antibodies, antibody receptors, cytokines and
XX chemokines, endogenously produced specific and non-specific enzymes,
XX binding proteins, adhesion molecules and their receptors, cytokine and
XX chemokine receptors, adenosine receptors, bradykinin receptors, central
XX nervous system (CNS) and peripheral nervous and non-nervous system
XX receptors, CNS and peripheral nervous and non-nervous system peptide
XX transmitters, defensins, growth factors, vasoactive peptides and
XX receptors, binding proteins and malignancy associated proteins. The
XX antisense oligonucleotides may be used in this way to treat disorders
XX including bronchoconstriction and/or lung inflammation, allergy(ies)
XX and/or surfactant hypoproduction which are associated with a disease or
XX condition selected from pulmonary vasoconstriction, inflammation,
XX allergies, asthma, impeded respiration, respiratory distress syndrome
XX (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
XX hyperextension, emphysema, chronic obstructive pulmonary disease (COPD),
XX pulmonary transplantation rejection, pulmonary infections, bronchitis,
XX and/or cancer. AAT18434 to AAT21543 represent human polynucleotide
XX fragments and antisense oligonucleotides used in the exemplification of
XX the present invention.

Oy	91	TTAAATATCAGAGATGTCATCTACACAGGACGGTTTACGTACACAAATTCATATGA	15
Db	333	TAGGAGATGGCCGGTGGCCACTGTAATGTCACACCGAAGAGCTCTG38CCAGCGCGAGGCC	392
Oy	151	GTTTCATATCCATACCGGAGACATACATAGTTGGATCTGTACAAATTTACCTACAGCAAGCTAG	210
Db	393	GTTGCTCTTCACACACCGGACACTGACAAACCCGGGCTCCCACTGCTCCCTG6CCACGG	452
Oy	211	CCACCCCTCCAGAAATGCAATTTGATATATTGGATTCACCAACTCAGACTCCCAAAAG	270
Db	453	CCGGAGGGGACCCCTGGGTGGGGCCCTCCATGGATCTCTAGCCACACCCAGGCTTC	512
Oy	271	AGTGGCATTTGAGCACTGTTTAACTCTGACCTTGACTTACCGCGCGTGAATCAGATATCCAA	330
Db	513	AGCCACCTCCAGAGCATCTTCACTGGGTGTCTAGCTACCGGCGGACTCGGACATCTTTT	572
Oy	331	GTCGCTTATGGCTCTTGTACGCGTAAAGCAAAATCCCTTGTTTGAAGTCCCAAGCAAA	390
Db	573	GTCGCTTATGGCGCGCTTGAGACCCCACTGGGGGCTCT---CCGCACGCTGCTCCAGCCAG	629
Oy	391	GAAATAATTTGGTGTGCTGGGTGTGAGTAACTGGAAACCTTGACATCCAGAGTCAGAT	450
Db	630	AGCAGAGGTGGCGGCTCGGTGGTGTACAGCAATCTTCAGAGACGGGCACTGCTGCGAGCTG	689
Oy	451	TACAAATGACCTTAAGCAAAAGCATTTGAATTCATACCTACAGGCGCAAGCAATTTGCAATAT	510
Db	690	TACCGGCACTGGCGCTCATCTCGGGGTGGATGTCTTTGGCGCGCAATGGAGCGGCCA	749
Oy	511	GTCATGATATTAATTTATTTATCTTACCATATATTCGTTGTAATTTTATCTTCTTTGAA	570
Db	750	CTGGGCGGCAAGTGGCTGTGCTCCACCGCTGGCCCACTACCGCTTCTACCTGCTCTTTGA	809
Oy	571	AATTCATATCCACAGAGATTACATATAGGAAACAT---ACAATGCTTTTCGGGCTGG	627
Db	810	AACCTTCACAGACCGGACTATACATTACGGAATAATTTGGCGGCAACGCACTGTGTGGCTG	869
Oy	628	TCTGTACGCTGTTCTTGTGGGACATCTTAGGAAACACTAGAAATATATATTCACAGAT	687
Db	870	ACTGTGCGAGTGTGCTGTGGGCCCCCACGGGCACTATAGAGGCTTGTGTGCTGGCTGAC	929
Oy	688	TCATTCATTCATGTGGAAAGATTATTACTCTCCAGTGAAGCTAGC	731
Db	930	GCCCTGTGCTCATGTGGATGTGACTTGTGCTACGCCGAGAGCTGGC	973
RESULT 9			
AAA35014			
ID AAA35014 standard; DNA; 1701 BP.			
AAA35014;			
28-JUL-2000 (first entry)			
Human adenosine receptor related polynucleotide SEQ ID NO:2703.			
Human: adenosine receptor; low adenosine antisense oligonucleotide;			
phosphorothioate; impaired respiration; inflammation; allergy;			
allergic disease; bronchoconstriction; inhibitor; antiinflammatory;			
anti-allergic; antiasthmatic; cytosolic; analgesic; impaired airway;			
lung disease; ischaemic condition; pulmonary vasoconstriction; asthma			
respiratory distress syndrome; pain; cystic fibrosis; emphysema;			
pulmonary hypertension; chronic obstructive pulmonary disease; COPD;			
cancer; leukaemia; lymphoma; carcinoma; metastasis; SS.			
XX Homo sapiens.			
OS WO200009525-A2.			
XX 24-FEB-2000.			
XX PD 99MO-US17712.			
XX PF 03-AUG-1999; 98DU-0095212.			
XX PR 03-AUG-1998; 98DU-0095212.			

XX (UYEC-) UNIV EAST CAROLINA.
 XX NYCE JW;
 XX MPI: 2000-205971/18.
 DR New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers
 XX
 PS Disclosure; Page 957-958; 1343pp; English.
 XX
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antisthmatic, cyostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing the
 CC bronchoconstriction and inflammation. AAA32313 to AAA33312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32313 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 CC
 XX Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;
 XX
 SQ
 Query Match 12.5%; Score 114.4; DB 21; Length 1701;
 Best Local Similarity 50.8%; Pred. No. 7.6e-23;
 Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

OY 511 GTCAATGATATAAAATTTGATTCCTACCATATCGCTTGTAATTTATCTTCCTTGAA 570
 DB 750 CTGTGGCGCAGCTGCTGGTGGCCACCGTGGCCAGTACGCTTACGCTCTTGAG 809
 OY 571 AATTCAATCCACAAGATTACATCAGCGAAACGTAI--ACAATGCTTTTGGCTGGC 627
 DB 810 AACTCTCAGCAGCCGCGCTCAATTACGAGAAATTCGTGGCCACACGACTGGGCTGGC 869
 OY 628 TCTGTACCTGTTGTTCTGGGACCATCTAGGAAACATGAGAAATATATATCCAGAGAT 687
 DB 870 ACTGTGCCAGTGTGCTGGGGCCCCACGGGCCACCTATGAGCCTTCTGTGGGCTGAC 929
 OY 688 TCATTCATTCATGGAAGATTATTAATCTTCCAGTGAGCTAGC 731
 DB 930 GCCTTGATGATGTGGATGACTTTGGCTCAGCCGAGAGCTGGC 973
 RESULT 10
 AAF21137
 ID AAF21137 standard; DNA; 6944 BP.
 XX AAF21137;
 AC
 AC 14-MAR-2001 (first entry)
 DE
 XX Human low adenosine antisense oligonucleotide related sequence #2704.
 XX
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antisthmatic; analgesic; hypotensive; cyostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RPS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200062736-A2.
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000MO-US08020.
 XX
 PR 06-APR-1999; 9905-0127958.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 PI NYCE JW;
 XX
 DR MPI: 2000-679539/66.
 XX
 PT Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions
 XX
 PS Disclosure; Page 1034-1036; 1592pp; English.
 XX
 CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antisthmatic, hypotensive and cyostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,

Db 5576 TAGGGATGCGCCGCGTCCACCTGAGTGGCCACCGAAAGCTGCTGGCCAGCGCCAGCGCC 5635
QY 151 GTTCTGATCCATACCGAGACATCACTGTTGGATCTGACAAATTTACCTCAGCAAGCTAGG 210
Db 5636 GTGGTCTTCCACCACCGCGAGCTGACAGACCCCGGGTCCACCTGCGCCCTGGCCAGCGG 5695
QY 211 CCACCCCTTCCAGAAATGATTTGGATGATTTGGAAATCACCACCACTACACTCCCAAAAG 270
Db 5696 CCGCGAGGCGAGCGCTGGGGTGGGCTCCATGAGTCTCTTCCAGCACACCCAGCGCTTC 5755
QY 271 AGTGGATGAGACACTGTTTAACTGACCTCTGACTTACCGCGGTGACATGATATCCAA 330
Db 5756 AGCCACTCCGAGGACATCTTCACTGGTGTGCTAGCTACCGCGGACCTGGACATCTTTT 5815
QY 331 GTGCTTATGGCTTCTTGAAGGTAGCACAATACTCCCTGCTGTTTGAAGTCCCAAGCAA 390
Db 5816 GTGCTTATGGCGCGCTGGAGGCCCACTGGGGCCCT---CGCCACCGCTGCCAGCCAAAG 5872
QY 391 GAGAAATTTGCTGTGCTGGTGTGATGACTGGAACCTGAGCATGCCAGATCCAAATAT 450
Db 5873 AGCAGGTTGGCCGCTGGGTGGTGCACCAACTTCCAGAGCGGACGCTGCTGCAGGCTG 5932
QY 451 TACATGAGCTAAGCAAAAGCATTTGAATTCATACCTACGCGGACGACATTTGGAGATAT 510
Db 5933 TACCGGACAGCTGGCGCTCATCTGCGGTGGATGCTTTGGCCGCTGCAATGAGACGCCA 5992
QY 511 GTCAATGATATAAAATTTGATCTTACCATATCTGCTGTAATTTATCTTCTCTTGA 570
Db 5993 CTGTGGCGCAGCTGCTGGTGGCCACCGCTGGGCCAGTACCGCTTCACTGCTCTTTGAG 6052
QY 571 AATTCAATCCACAAGATTTACATCAGCGAAAGCTAT---ACAATGCTTTTGGCTGGC 627
Db 6053 AACTCTCAGCAGCGCACATTAAGGAGAAATTTCTGGCGCAACGACACTGGTGGCTGGC 6112
QY 628 TCTGACCTGTGTGCTGGGACCATCTAGGAAACTATGAGAAATTAATTATTCAGAGCAT 687
Db 6113 ACGTGTGCACTGTGCTGGGGGCCCCCAGCGGCCACCTATGAGGCTCTTGGCCGGCTGAC 6172
QY 688 TCATTCATTGATGAGAGATTAATTAATCTCCAGTAGAGCTAGC 731
Db 6173 GCCTTGATGATGTGATGACTTTGGCTCAGCCGAGAGCTGGC 6216

RESULT 12
AAQ56905
ID AAQ56905 standard; DNA; 1654 BP.
XX
AC AAQ56905;
XX
DT 26-JUL-1994 (first entry)
XX
DE PCDNA1-alpha-(1-3)Fuc-TV1 genomic DNA.
XX
KM Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
XX surface; oligosaccharide; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 128..1208
FT /tag= a
XX
PN W09402616-A.
XX
PD 03-FEB-1994.
XX
PF 20-JUL-1993; 93WO-US06703.
XX
PR 20-JUL-1992; 92US-0914281.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Lowe JB;

XX WPI; 1994-048874/06.
DR P-PSDB; AAR45933.
XX
PT DNA fragment encoding a glycosyltransferase - can be used for in
PT vitro reactions to modify cell surface oligosaccharide(s) e.g.
PT blood gp. determinants, to protect against transplant rejection
XX
PS Claim 3; Fig 7; 249pp; English.
XX
CC The sequence is that encoding human GDP-Fuc: [beta-D-Gal (1-4)]-D-
CC GlucNAc alpha (1-3)-fucosyltransferase. The enzyme produced by the
CC DNA may be non glycosylated. This prevents premature loss of enzyme
CC activity. It can also be used in in vitro reactions to modify cell
CC surface oligosaccharide mols. e.g. blood group determinants.
CC See also AAQ56906-12.
XX
SQ Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;
Query Match 12.4%; Score 113.4; DB 15; Length 1654;
Best Local Similarity 51.2%; Pred. No. 1.5e-22;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
QY 31 ATTCTGTGTGGGTGGCCATTTGGGACAGCTTTGACCTTACATCTCCCAAGCATG 90
Db 318 ATCTGCTGTGGAGCTGGGCTTTTAAACAAACCATAGCTCTGCCCCGCTCAGAGATG 377
QY 91 TT---CAACATCCAGAGATGCCATCTCACAGGAGCGGTCACTGTACAAACAAATCCCAT 147
Db 378 GTGCTTGGCACGGCTACTACTGCAATCATCTGCGGACCGCAAGGTGTATCCACAGCAGAC 437
QY 148 GCAGTTCTGATCCATCAGCAGACATCACTGTTGGATCTGA---CAATTTACTCTCAGCA 204
Db 438 GCGGTATCGTGCACACCGAGAGGTGATGTACAAACCCAGTCCCAAGCTCCACAGCTCC 497
QY 205 GCTAGGCCACCTTCCAGAAATGATTTGATGAATTTGGAAATACCAACTCACTGCC 264
Db 498 CCGAGGCGGAGGAGGAGGAGGATGATCTGTTGAGATGAGTCCCAAGCCACTGCTGG 557
QY 265 CAAGAAGTGATGAGCACTGTTTAACTGACTCTGCTGCTGACTACCGCGTGTATGAT 324
Db 558 CAGCTCAAAAGCCATGACGAGATCTTCAATCTACCATGCTCTTACCGACGCGCTCCGAC 617
QY 325 ATCCAAGTGCCTTATGCT-----TCTTGACGCTAAGCAAAATCCCTTCGTTT 375
Db 618 ATCTTCACGCTTACAGGCTGGTGGAGCCGTGGTCCGGCCAGCTCCACCCAGCGCTC 677
QY 376 GAAGTGCACAGCAAGAGAAATTTGTTGCTGGCTTTGTGAGTACGTGAACCTGAGCAT 435
Db 678 AACCTCTCGGCCAAGACCGGCTGGCTGCGCAGTGTCCAACTGGGGGCCAAACTCC 737
QY 436 GCCAGAGTCAGATTAATCAATGAGCTAAGCAAAAGATTAATCAATCACTACCTGCGGCA 495
Db 738 GCCAGGCTGCGCTACTACCAAGACCTGAGGCCCACTCAAGGTGACGCTGAGCG--A 794
QY 496 GCATTTGAGAAATATGCTCAATGATTAATAAATTTGATCTACCATATCTGCTGTAATTT 555
Db 795 CGCTCCACAAAGCCCTGCCCCAGGAGAACCATATGAGACGCTGTCGGGTCAAGTTCC 854
QY 556 TATCTTTCTTTAAATTAATCAATCCACAGAGATTAATCACTGCGAAAGCTAT---ACAAT 612
Db 855 TATCTGCGCTTGAAGATCTTCAACCCGACCCGACTATACCCAGAGAGCTGTGAGAGAC 914
QY 613 GCTTTCTGGCTGGCTCTGACTGTTTCTGAGGACATCTGAGAAACATATGAGAT 672
Db 915 GCCCTGAGAGGCTGGGCGCTGCGCTGTGGTGGCCCAAGCAAGCAATGAGAGG 974
QY 673 TATATTCAGACAGATTCATTCATCTGAGAGATTAATTAATCTCCAGTGAAGCTAGCA 732
Db 975 TTCTGCGACCGAGCGCTTCAATCAAGTGAAGACTTCCAGAGGCCCAAGAGACTGGGC 1034
QY 733 AAGTATCTGAAGAGATCGACAAACAAATAGTTTATCTTATGTTTAACTGAGAG 792

Db 1035 CGGTACCTGACAGAGCTGGACAGACCCCGCTACCTGACCTTTCCGTGGCG 1094
QY 793 AAG 795
Db 1095 GAG 1097

RESULT 13

AAT61680
ID AAT61680 standard; DNA: 1654 BP.

AC AAT61680;

DT 19-JUN-1997 (first entry)

DE Human alpha(1,3)-fucosyltransferase (Fuc-TVI) DNA.

KW Alpha(1,3/1,4)-fucosyltransferase; Lewis enzyme; Fuc-TVI;
glycosylation; oligosaccharide; blood group; ss.

OS Homo sapiens.

FI Key Location/Qualifiers
FT CDS 129..1208
ET /*tag- a

PN WO9709421-A1.

PD 13-MAR-1997.

PF 06-SEP-1996; 96WO-US13816.

PR 08-SEP-1995; 95US-0525058.

PA (UNMI) UNIV MICHIGAN.

PI Legault DJ, Lowe JB;

DR WPI, 1997-192897/17.

DR P-FSDB; AAM13643.

PT New recombinant fucosyltransferase proteins - useful for modifying
cell surface oligosaccharide structures

PS Example 6; Page 291-292; 329pp; English.

CC A DNA clone (AAT61680) codes for an enzyme (AAM13643) that functions as
a GDP-Fuc:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-fucosyltransferase
(Fuc-TVI) which can utilize neutral type II oligosaccharide
precursors but not type I glycoconjugates. It was identified in a
human lambda phage genomic DNA library using a probe obtd. from
CC Lewis enzyme cDNA (see also AAT61675). The Fuc-TVI DNA can be used
to construct animal cell lines with specific capabilities with
CC respect to post-translational modification of the oligosaccharides
of expressed proteins or lipids, to produce recombinant Fuc-TVI for
use in oligosaccharide prodn., to genotype individuals at this
CC fucosyltransferase locus, and in the prodn. of chimera
CC fucosyltransferases (see also AAM14512-32).

XX Sequence 1654 BP: 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match 12.4%; Score 113.4; DB 18; Length 1654;

Best Local Similarity 51.2%; Pred. No. 1.5e-22;

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

QY 31 ATTCTGCTGGTGGTGGCTATTGGCAGACCTTGACCTTACATCCCAAGCAATG 90
DB 318 ATCTGCTGTGACCTGCTTTTAAACAACCCATAGCTCTGCCCTGCTCAGAGATG 377
QY 91 TT---CAACATCCAAAGTGCATTCACAAAGAGAGCTTCACAAACAAATCCAT 147.
DB 378 GTGCTGACAGCGCTGATGACATCATCTGCCAGCCCAAGGTGTATCCACAGGAGAC 437

QY 148 GCAGTTCTGATTCATCACCCAGACATCATGTTGGATCTCA---CAAAATTACCTCAGCA 204
DB 438 GGGTCATCTGTGACACACCGAGAGTATGATACAAACCCAGTGCACAGCTCC 497
QY 205 GCTAGGCCACCTTCCAGAAAATGGATTGGATGAAATTTGGAATCACCACACTCC 264
DB 498 CCGAGCGCGGAGGGGACCGATGATCTGTTGACATGAGATGCCCAACCACTGCTCG 557
QY 265 CAAAAGAGTGGCATTTGAGCACTTTTAACTGACTCTGACTTACCTGAGCTTACAT 324
DB 558 CAGCTGAAGAAGCATGAGACGATACCTCAATCTCACCATGTCTTACCGACGACTCCGAC 617
QY 325 ATCCAAAGTCCCTTATAGCT-----TCCTGACGATGAGCAAAATCCCTGCTGTTT 375
DB 618 ATCTTCACAGCCCTACGGCTGAGCGCTGGACCGTGGTCCGCGGACGCTGCCACACCGCTC 677
QY 376 GAAATGCCCAAGCAAGAAATTTGTTGCTGCTGCTGTTGAGTAACTGCAACCTGAGCAT 435
DB 678 AACCTCTGGCCCAAGACCGAGCTGTGCTGGCAGTGTCCACATGAGGCGCAAACTCC 737
QY 436 GCCAGATCAAGTATTTACATGAGCTAAGCAAAAACATTTGAATTCCTACGGGCA 495
DB 738 GCCAGGATGCGCTACTACACAGAGCTGACAGCCCATCTCAAGGTGAGCAGTGAACG---A 794
QY 496 GCATTGGAGAAATATGTCATGATMAAAATTTGATTCCTACCATATCTGCTGTAATTT 555
DB 795 GCTCCCAAGACCCCTGCCCCAGGGAACCATGATGAGACGCTGTCGCGTACAAAGTTC 854
QY 556 TATCTTCTTGAATAATTCATCCACAGAGATTACATCCAGGAAGTAT---ACAT 612
DB 855 TATCTGCGCTTGAAGAACTCTTCCACCGGACTATACATCCAGAGATGTGGAGAAC 914
QY 613 GCTTTCTGCTGCTGCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
DB 915 GCTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974
QY 673 TATATTCAGCAGATTTATTCATTCATGATGAGATTAATTAATCTCCACAGTACGA 732
DB 975 TTCTGCGCCACCGAGCGCTTATTCACGTGACGACTCCAGAGCCCAAGACCTGGCC 1034
QY 733 AAGTATCTGAAGGAGTGCACAAAACAAATTAATTAATCTTAACTTAACTGAGG 792
DB 1035 CGTACCTGACGAGCTGACAAAGACACGCGCTACCTGACTGACTTTCGTGGCGG 1094
QY 793 AAG 795
DB 1095 GAG 1097

RESULT 14
ID AAO56912
AAO56912 standard; DNA: 1086 BP.

AC AAO56912;

DT 26-JUL-1994 (first entry)

DE Lewis blood group fucosyltransferase DNA.

KW Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;

XX surface; oligosaccharide; ss.

PN WO9402616-A.

PD 03-FEB-1994.

PF 20-JUL-1993; 93WO-US06703.

PR 20-JUL-1992; 92US-0914281.

PA (UNMI) UNIV MICHIGAN.

PI Lowe JB;

```

XX  WPI: 1994-048874/06.
XX
XX  DNA fragment encoding a glycosyltransferase - can be used for in
XX  vitro reactions to modify cell surface oligosaccharides) e.g.
XX  blood gp. determinants, to protect against transplant rejection
XX
XX  Disclosure; Fig 6; 249pp; English.
XX
XX  The sequence is that encoding a Lewis blood group fucosyl
XX  transferase (Fuc-TIII). The enzyme produced by the DNA may be
XX  non glycosylated. This prevents premature loss of enzyme activity.
XX  It can also be used in in vitro reactions to modify cell surface
XX  oligosaccharide moles. e.g. blood group determinants.
XX  See also AA056905-11.
XX
SQ  Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;

Query Match      12.3%; Score 112.2; DB 15; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.7e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY  106 TGGCATCTCACAAGCGGTTTCACCTGATACCAACAATCCCATGAGTTCGATCCATCAC 165
DB  271 TGGCATATCATCTGCCACCGCAAGGTGTACCCACAGGACACACGGTCTATGTCACACC 330
OY  166 CGAGACATCAGTTGGGA---TCTGACAAATTTTACCTCAGCAAGCTAGGCGCACCTTTCAG 222
DB  331 TGGGATATATGTCACCAACCTTAGTCACGGCTCCACCTTCCCGAGGCCGCGAGGGGCGAG 390
OY  223 AATGATTTGGATGATTTGGAAATGACCACTACACATCTCCCAAAAAGAGTGCATTGAG 282
DB  391 CGCTGATCTGTTCAACTTGGAGCCACCCCTTACCTGTCAGACCTGGAAAGCTCTGGAGC 450
OY  283 CACTGTTTAACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 342
DB  451 AGTACTTCAATCTCACCATCTGCTCTCCGACGAGCTCCGACATCTTCCACGCCCTTACGG 510
OY  343 TTCTTGACGGTAAACCAATCCCTTCTGT-----GTTTGAAGTGGCCAAAGCAAGAG 393
DB  511 TGGCTGAGAGCGGTGTCGGGCCAGCGCCACCCACCGCTCAACCTCTCCGCGCAAGAC 570
OY  394 AATTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 453
DB  571 GAGCTGTGGCTGGCGGGGCTGTCCAACTGAAACCGGACTCAGCAGGCTGCGCTACTAC 630
OY  454 AATGAGCTAAAGCAAAAGCATTTGAAATCCATACCTACGGGGCAAGCTTGGAGAAATATGTC 513
DB  631 CAGAGCTGCAAGGCTCATCTCAAGGTGAGCTGTAGGAGCGCTCCACAAAGCCCTGCCCC 690
OY  514 AATGATTAATAATTTGATCTCTACCATATCTGCTTGAATTTTATCTTCTTGAATAAT 573
DB  691 AAGGGAGACA---TGATGAGAGAGCGTGTCCCGGTACAAAGTTTACCTGGCGTTTGAAGAC 747
OY  574 TCAATCCACAAGATTTACATCAGGAAAAAGCTAT---ACAAGCTTTTCTGGCTGGCTCT 630
DB  748 TCTCTTCAACCCGACACATCACTACACGAGAACTGTGAGAGAACCCCTGGAGGCTGGGCC 807
OY  631 GTACCTGTTGTTCTGGGACCATCTAGGAAAACTATGAAATTAATTCACAGAGATTTCA 690
DB  808 GTGCCCGGTGTGTGGGCCCCCAGACAAAGCAACTACGAGAGGTCTCTGCCACCCGACGCC 867
OY  691 TTTCATTGATGTGAGAGATTATATACTCTCCAGTGAAGCTAGCAAAATATCTGAAGGAAGTC 750
DB  868 TTTCATCCAGTGTGAGAGACTTCCAGAGCCCAAGAGACCTGGCCCGGTACCTGCGAGGAGCTG 927
OY  751 GACAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 795
DB  928 GACAAGAGCCACGCGCGCTACCTGAGACTTCTGCTGGCGGAG 972

```

```

ID  AAT80111 standard; cDNA; 2042 BP.
XX
XX  AAT80111;
AC  AAT80111;
XX
XX  01-DEC-1997 (first entry)
DT
XX
XX  Fucosyltransferase III coding sequence.
DE
XX
XX  P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;
XX  calcium-dependent carbohydrate binding protein; platelet; endothelium;
XX  thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;
XX  inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
XX  disseminated intravascular coagulation; rheumatoid arthritis; antibody;
XX  adult respiratory distress syndrome; tumour metastasis;
XX  fucosyltransferase III; acetylglucosaminyltransferase;
XX  autoimmune disorder; inflammatory disorder; ss.
XX
OS  Homo sapiens.
XX
XX  W09706176-A2.
PN
XX
XX  20-FEB-1997.
PD
XX
XX  02-AUG-1996; 96MO-US12820.
PP
XX
XX  17-MAY-1996; 96GS-0649802.
PR  03-AUG-1995; 95US-0510920.
PR  15-MAY-1996; 96US-0017794.
XX
XX  (OKLA ) UNIV OKLAHOMA STATE.
PA
XX
XX  Cummings RD, McEever RP, Moore KL;
PI
XX
XX  WPI: 1997-154206/14.
DR
XX
XX  New O-glycan(s) derived from P-selectin glyco:protein ligand
XX  (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
XX  for treating inflammatory or autoimmune disorders or tumours
XX
XX  Disclosure; Page 80-81; 96pp; English.
XX
XX  This sequence represents the coding sequence for the human
XX  fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-
XX  acetylglucosaminyltransferase encoded by AAT80112 are needed for the
XX  expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in
XX  mammalian cells. P-selectin is a calcium-dependent carbohydrate binding
XX  protein expressed on the surfaces of activated platelets and endothelium
XX  in response to thrombin and other agonists. PSGL-1 (see AA056174) is a
XX  high affinity P-selectin ligand produced by leukocytes. Binding of
XX  P-selectin to PSGL-1 is calcium ion dependent and is abolished by
XX  treatment of the ligand with sialidase. PSGL is a homodimer, with a
XX  highly extended extracellular domain, which is a feature of mucin-like
XX  proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of
XX  the invention (see AA056175-W26188). PSGL-1 is also targeted by the
XX  O-glycans, and O-glycanated products of the invention. The products are
XX  used for inhibiting the binding of P-selectin and other selectins to
XX  PSGL-1. They can be used for the treatment of inflammation, injury
XX  resulting from ischaemia and reperfusion, bacterial sepsis and
XX  disseminated intravascular coagulation, adult respiratory distress
XX  syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and
XX  other autoimmune or inflammatory disorders. The products can also be used
XX  in the production of antibodies and in diagnostic applications.
XX
SQ  Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

Query Match      12.3%; Score 112.2; DB 18; Length 2042;
Best Local Similarity 51.6%; Pred. No. 3.6e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY  106 TGGCATCTCACAAGCGGTTTCACCTGATACCAACAATCCCATGAGTTCGATCCATCAC 165
DB  343 TGGCATATCATCTGCCACCGCAAGGTGTACCCACAGGACACACGGTCTATGTCACACC 402

```

RESULT 15
AAT80111

Tue Oct 8 10:18:29 2002

us-09-744-748-4_copy_1454_2368.rng

Page 14

OY	166	CGAGCACTACACTTGGGA---	-TCGACAATTTACTCCAGCAAGGTATGAGGCCACCTTCAG	222
Dd	403	TGGGATCATCATGTCACACCCTTAAGTCAGAGCCTCCACACTTCCCAGAGCCGCAGAGGGCAG	462	
OY	223	AAATGATTGGTAGAATTTGGATATCACACACACACACCTCCCCAAAAGAATGACATTTAG	282	
Dd	463	CCTTGATCTCGTTCACTTTGGAGCCACCCCTTACTCTGACAGCACTTGAAGCCCTGGAC	522	
OY	283	CACCTGTTAACTGACTCTGACTTACCGCCGTGATTCAGATATCCAGTGGCTTATGSC	342	
Dd	523	AGATCTTCATATCTACCATTTCTCTACCCGACGAGCACTCCGACATCTTCAAGCCCTTACGGC	582	
OY	343	TTCTTGACGGTAAGCACAAATCCCTTGST-----GTTTGAAGTGCACCAACAAGAG	393	
Dd	583	TGGCTGGAGCCGTGATCCGGCACCTGCCCCACACCCGCTCAACCTCTCGGCCAGAAC	642	
OY	394	AAATGGATGCTCGGTGTGAGTAACTCGGAACCTCGAGCATGCGACAGTCAAGTATTAC	453	
Dd	643	GAGCTGGTGGGCTGGGGGGGTGTCCAACITGGAAAGCCGAGACTCAGACAGGTGGCCCTACTAC	702	
OY	454	AATGAGCTAACCAAAAGATTTGAATTCATCTACGGSGCAGCATTTGGAGAAATATGTC	513	
Dd	703	CAGAGCTGCAAGGCTCATCTCAAGATGGACGTGTAGGAGAGCTTCCACAAGGCCCTGCCCC	762	
OY	514	AATGATAAAAATTTGATTCCTACCATATCTGCTGTATAATTTATATCTTCCCTTGAAGAT	573	
Dd	763	AAGGGGACCA--TGATGGAGAGAGCTCTCCCGGTACAAAGTTCTTACCTGGGCTTCGAGAAC	819	
OY	574	TCAATCCACAAAGGATTACATCAGCGAAAAAGCTAT---ACATAGCTTTCTGGCTGGCTCT	630	
Dd	820	TCTCTTGACCCCGACTACATACCGAAGAGCTGTGAGAGAAACCCCTCGAGGCTGGGCC	879	
OY	631	GTACCTGTTGTTTGGGACCACTATAGGAAAAACATATAGAAATTAATATATTCAGAGATTCA	690	
Dd	880	GTGCCCGGTGGTGGGGCCCCAGAGAAGCAACTACGAGAGGTTCCTGCGCACCCGAGGCC	939	
OY	691	TTTCATTCATGTGGAAAGATTAACTCTCCAGTAGGTATACAAAGTATTCGAAGGAAGTC	750	
Dd	940	TTTCATTCACGTGGAGCACTTCCAGAGCCCAAGAGACCTGGGCCGGTACCTGCGAGAGCTG	999	
OY	751	GACAAAAACATTAAGTATATACCTTAGTACTTTTACTTGAAGAGAG	795	
Dd	1000	GACACGAGACACAGCCCGCTACTAGTACTTCTGCTGGCGGGAG	1044	

```
Search completed: October 5, 2002, 23:13:52
Job time : 155.682 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 766.105 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-4_COPY_1454_2368

Perfect score: 915
Sequence: 1 accaaactgattatttaa.....agaatggtttgaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	573.8	62.7	697	9	BB397507 BB397507
2	492.8	53.9	607	10	B1102760 602888338
3	422	46.1	716	9	AV725071 AV725071
4	324.8	35.5	649	9	AV338075 AV338075
5	307.4	33.6	579	10	BM021081 BM021081
6	286	20.1	1101	12	CNS059ST CNS059ST
7	184.2	20.1	570	10	BM082484 BM082484
8	177	19.3	1100	12	CNS05792 CNS05792
9	170.2	18.6	1016	12	CNS0429P CNS0429P
10	161.4	17.6	844	12	CNS018V CNS018V
11	157.4	17.2	887	12	CNS020S CNS020S
12	155.4	17.0	651	10	Bj036855 Bj036855
13	151.6	16.6	666	10	Bj050300 Bj050300
14	151.2	16.5	711	12	CNS034XF CNS034XF
15	151	16.5	501	12	BA1073 BA1073
16	146.6	16.0	640	10	Bj073227 Bj073227
17	145.2	15.9	648	9	BB595164 BB595164

18	136.2	14.9	199	9	BB398103 BB398103
19	133	14.5	625	9	BB649967 BB649967
20	123.8	13.5	767	9	BB640911 BB640911
21	123.4	13.5	674	10	BB605503 BB605503
22	120.6	13.2	971	12	CNS0471V CNS0471V
23	120.4	13.2	615	10	B1326811 B1326811
24	117.8	12.9	1023	12	CNS02ANG CNS02ANG
25	102.8	11.2	487	10	BB683396 BB683396
26	100.2	11.0	702	10	Bj043578 Bj043578
27	93.8	10.3	863	12	CNS02ROF CNS02ROF
28	90.4	9.9	720	10	BF180380 BF180380
29	90	9.8	551	10	Bj035701 Bj035701
30	87.4	9.6	601	10	Bj036679 Bj036679
31	87.4	9.6	605	10	Bj028804 Bj028804
32	87	9.5	1056	12	CNS01VU7 CNS01VU7
33	86.8	9.5	334	9	AU081478 AU081478
34	84.2	9.2	638	9	Bj035697 Bj035697
35	83.6	9.1	1057	12	CNS03ADC CNS03ADC
36	83	9.1	566	10	Bj090565 Bj090565
37	80.2	8.8	628	10	BB651409 BB651409
38	79.4	8.7	776	12	CNS03R10 CNS03R10
39	78	8.5	568	10	BM021317 BM021317
40	78	8.5	838	10	BG432556 BG432556
41	76.4	8.3	655	10	BG338740 BG338740
42	75	8.2	655	10	CNS02V1D CNS02V1D
43	71.6	7.8	590	10	BM082755 BM082755
44	70.4	7.7	582	9	AM233458 AM233458
45	70	7.7	505	10	Bj029303 Bj029303

ALIGNMENTS

RESULT 1
LOCUS BB397507 697 bp mRNA linear EST 24-OCT-2001
DEFINITION BB397507 RIKEN full-length enriched, ES cells Mus musculus cDNA
clone C330004C19 3' similar to AB015426 Mus musculus Fcγ mRNA for
αpha1.3-fucosyltransferase IX, mRNA sequence.

ACCESSION BB397507.2 GI:16411808
VERSION BB397507
KEYWORDS
TERMS

SOURCE
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 697)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyata,T.,
Muramatsu,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)

TITLE
JOURNAL
COMMENT
On Jul 15, 2000 this sequence version replaced gi:9216903.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagl,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Matahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura


```

/tissue-type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies) . cDNA
made by oligo-dt priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
BASE COUNT      203 a      93 g      180 t
ORIGIN

```

Query Match	33.6%;	Score 307.4;	DB 10;	Length 579;
Best Local Similarity	99.7%;	Pred. No. 2.6e-74;		
Matches 308;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 607	TACAATGCTTTTTCGGCTGCTCTGTACCTCTTGTTCGGAGCACCATCAGGAAACTAT	666		
Db 579	TACAATGCTTTTTCGGCTGCTCTGTACCTCTTGTTCGGAGCACCATCAGGAAACTAT	520		
QY 667	GAGAAATTAATTCACGACGATTCATTCTATGTGGAAGATTATTAAGTCTCCAGTGA	726		
Db 519	GAGAAATTAATTCACGACGATTCATTCTATGTGGAAGATTATTAAGTCTCCAGTGA	460		
QY 727	CTACGAAGTATTCGAGGAGAGTCGACAAAAACAATAGTAATACCTAGTACTTTTAC	786		
Db 459	CTACGAAGTATTCGAGGAGAGTCGACAAAAACAATAGTAATACCTAGTACTTTTAC	400		
QY 787	TGAGGAGAGGATTTCACTGTAAATCTTCCACGATTTTGGGATCAGATGTTTGCT	846		
Db 399	TGAGGAGAGGATTTCACTGTAAATCTTCCACGATTTTGGGATCAGATGTTTGCT	340		
QY 847	TGCGATCATGTGAAAGGCATCAAGATATAAGTCTGTTGTAATTAGAGAAATGTTTT	906		
Db 339	TGCGATCATGTGAAAGGCATCAAGATATAAGTCTGTTGTAATTAGAGAAATGTTTT	280		
QY 907	TGGAATTAA 915			
Db 279	TGGAATTAA 271			
RESULT 6				
CNS059ST/c				
LOCUS				
DEFINITION	CNS059ST	1101 bp	DNA	linear
	Tetradon nigroviridis genome survey sequence T3 end of clone			
	045P22 of library A from Tetradon nigroviridis, genomic survey			
	sequence.			
ACCESSION	AL327638	1	GI:8221227	
VERSION	AL327638			
KEYWORDS	GSS; genome survey sequence.			
SOURCE	Tetradon nigroviridis.			
ORGANISM	Tetradon nigroviridis			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;			
	Tetraodontidae; Tetradon.			
REFERENCE	1 (bases 1 to 1101)			
AUTHORS	Roest-Crolius,H., Jallion,O., Dasilva,C., Fiazmes,C., Fisher,C.,			
	Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and			
	Weissenbach,J.			
TITLE	Characterization and repeat analysis of the compact genome of the			
	freshwater pufferfish Tetradon nigroviridis			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1101)			
AUTHORS	Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,			

TITLE	Human gene number estimate provided by genome wide analysis using Tetradodon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1101)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetradodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetradodon .
FEATURES	Location/Qualifiers
source	1..1101

```

/db.xref="taxon:99883"
/clone="045P22"
/clone.lib="A"
/note="Genoscope sequence ID : COAA045DH1AI-end : "T3"
BASE COUNT      287 a      249 c      269 g      278 t      18 others
ORIGIN
Query Match      31.3% Score 286; DB 12; Length 1101;
Best Local Similarity 59.1%; Pred. No. 2.7e-68;
Matches 537; Conservative 0; Mismatches 366; Indels 5; Gaps 3;

QY   12 TTATTTTAATGAAMACTACTATTCTGGTGTTGGGTGAGCCATTTTGGGCAGACCTTTGACCT 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   941 TGATAAAAACAATGACCATTTATCTCTGTGGTTGGCGCCTTCGGCAACAACCTTGACCT 882

QY   72 TACATCCCGCAAGAACAATGTTCAACATCCAGAGATGCCATCTCACAAAGGACCGCTTACC 131
    | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   881 GAACGTGTGCAACGGTTATTTTCGGTTTGAGGGCTGTTCACATTTACGGCGGACAGAAACT 822

QY   132 GTACACAAAATCCCATCTGATTTTCATCTCAGCAGACATCACTWTGGATGTACAAA 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   821 TTTTAACAAGTCGATGGCGTTGTCTTCACACAGAACATCGCCGGGTGACCTTCCA 762

QY   192 TTTACCTCAGCAAGCTAGGCCACCTTCAGAAATGATTTGGATGAATTTTGAATACCC 251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   761 CATGCCGACAGTCGTGGCCACCGCTCCGATCAAAAATGATCGTTTNNCATGAGACTTCC 702

QY   252 AACTCAGACCTCCCCAAAAGATGGGATTAGAGACACTGTTTAACCGTCACTGACTTACCG 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   701 CTCACACTCTCTTAGGCACCTGTGATGTAGAGGCTGTTCACCTGACGCTGAACTATCG 642

QY   312 CCGTGAATCAGATATCCAAAGTGCCCTTATGGCTTCTTGAAGGTAAGCAAAAATCCCCTTGT 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   641 TCCTAGATTCGACATTTTCCACGTCGCTTTATTTATCCCTGTGTGCACTGATGCTGAAGA 582

QY   372 G---TTTAAATGCCAGCAAGAGAATAATGGTGTGCTGGGTTGTGAATTAACGTGAACCC 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   581 GGACTTTGTCCACCTTACCAAAAACAAAACCTGTTGGTGGATTTGTCACAACTGTGAACCA 522

QY   429 TGAGCATCCAGAGTCACAGTATTACAATGAGCTAAGCAAAAAGCATTTGAATTCATACCTA 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   521 GGAAATCTCAGAGATGAATTTCTACAAAGATTTGTACAGAGATGTTGTAGGTGCCAACGTA 462

QY   489 CGGGCAAGCAATTTGGAGAAATATGTCAATGATPAATAATTTGATTCTCAATATCTGCTTGG 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   461 TGGAGAGACATTTGGGGCAACACTTCCCGAGATGAAGAGACCTTACCAACACATCGCAAGTG 402

QY   549 TAAATTTATTTCTTTCCTTTGAAATTCAAATCCAAAGAGATTAACATCAAGGAAAGACTATA 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   401 CAAGTTCTACCTGTGATTTTGAAGAACTCCATCCACAAAGACTACTTTTACTGAAAACTGTA 342

QY   609 CAATGCTTTTCTGGCTGCTCTGTACCTGTTGTCTTGAGGACATCTAGAGAAACTATAGA 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   341 CAACCACTTCTGTGGGGACAGTGGCCAGTGTGTTCTGGGGCCACCTCGGGAAGACTAGA 282

QY   669 GAATTAATTTCCAGCAAGTTCAATTCATTCATCATGTGGAAGATTATAACTCTCCCAAGTACCT 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 281 GACATCATCCAGGAGATGCTTTATCCAGTGGAGACTTGGCTTCAGCCAGAGACT 222
 Oy 729 AGCAAGATCTGAGAGAGAGTGCACAAAACATTAAGTATACCTTACCTTAACTG 788
 Db 221 TGCTGACTATTTACTGCTCTGCAAAAATGAGAAATGACTTAAATATTTTCAAGT 162
 Oy 789 GAGAGAGATTTGACTGTAATCTCCAGATTTTGGGAATCAACATGATGTGGCTG 848
 Db 161 GCGCAAGACACTTAAAGCAAGGAGTGCATTTCCCGAGCC-CACCTGACAGCTTG 103
 Oy 849 CGATCAGTGAAGGATCAAGATATA-AGTCCTGTGTGTAATTTAAGAAAGCTTT 907
 Db 102 TGATTACGTCAGAGGAGTANCAATACAGATTCAGTGCATTANCAAGTGTACT 43
 Oy 908 GGAATTA 915
 Db 42 GGGGTGA 35

RESULT 7
 BM082484/c 570 bp mRNA linear EST 14-NOV-2001
 LOCUS fu25608.x1 Campbell zebrafish ovary Danio rerio cDNA clone 5306223
 DEFINITION 3' similar to TR:088819 088819 ALPHAI.3-FUCOSYLTRANSFERASE IX. ;
 mRNA sequence.

ACCESSION BM082484
 VERSION BM082484.1 GI:16929414
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.
 1 (bases 1 to 570)
 Clark,M., Johnson,S.D., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,
 S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Underwood
 R., Steptoe,M., Theisling,B., Allen,M., Bowers,Y., Person,B.,
 Waller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R., Ritter,E.,
 Kohn,S., Shiu,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
 and Wilson,R.
 Mashu zebrafish EST Project 1998
 Unpublished (1998)
 Other ESTs: fu25608.y1
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbralis@wustl.edu

TITLE JOURNAL
 COMMENT

CNA Library constructed by Library constructed by Invitrogen and
 donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA).
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 High quality sequence stop: 419.
 Location/Qualifiers
 1..570
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="5306223"
 /clone_1ib="Campbell zebrafish ovary"
 /sex="female"
 /dev_stage="4-5 months, 1 year and 2 years"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary(pooled); Vector: PCMV-SpOrf6, Site_1:
 NotI; Site_2: SalI; Whole ovaries collected from zebrafish
 aged 4-5 months, 1 year and 2 years. Oligo-dT primed,
 directionally cloned. Average insert size 2 kb. Library
 constructed by Invitrogen and donated by R. Campbell
 (Marine Biology Laboratory, Woods Hole, MA)."

FEATURES

source

BASE COUNT
 ORIGIN

155 a 106 c 129 g 180 t

Query Match 20.1%; Score 184.2; DB 10; Length 570;
 Best Local Similarity 59.2%; Pred. No. 3, 6e-40;
 Matches 337; Conservative 0; Mismatches 223; Indels 9; Gaps 1;

Oy 217 TTCGGAAGATGGATTTGGATGATTTGGATTCACCAACATCCACACCCCAAGAGTGGC 276
 Db 570 TTCGGAAGATGGATTTGGATGATTTGGATTCACCAACATCCACACCCCAAGAGTGGC 511
 Oy 277 ATTAGACACTGTTTAACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 336
 Db 510 CTGGAAACCTGTTCAACACACATCTGATCTGATCTGATCTGATCTGATCTGATCTG 451
 Oy 337 TATGCTCTTGAAGGTAGCAAAATCCCTGCTGTTGAAGTGCAGCAAGAGAGAA 396
 Db 450 ATGAGATTTAAACACAGAGAAACAGCCGATGATTTGTCTATCCCAAGAGAGAA 391
 Oy 397 TTGCTGTGCTGGCTTGTGAGTAACTGGAACCC-----TGACATGCGAGACTCAAG 447
 Db 390 CTAGTTTGTGATGTGAGTAACTGGAATTCCTCCATTAACAGAGTGTGAGCAAGAGAG 331
 Oy 448 TATTCAATGAGCTAAGCAAAAGCAATTAATCCATACCTGAGGCAAGCAATTTGGAGAA 507
 Db 330 ATTATGCGGAATTAAGCGCATACATCATGATACACCTTTGCGGAGGCTTATTCACAA 271
 Oy 508 TATGCAATGATTAATAATTTATTCCTACCAATATCTGCTTGAATTTATCTTCTT 567
 Db 270 TATCTAGACTATTAAGACTACTACTGCTGCTGCAAAATTTTACCTCTCTTC 211
 Oy 568 GAAATTCATCCACAGAGATTTACATGAGGAAAGCTATACATGCTTTCTGCTGCGC 627
 Db 210 GAAATTCATCCACAGAGATTTACATGAGGAAAGCTATACATGCTTTCTGCTGCGC 151
 Oy 628 TCTGACCTGTTGTTCTGCGACATCTAGGAAAGCAATATATTTCCACAGAT 687
 Db 150 ACTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 91
 Oy 688 TATTCATCTGATGAGAGATTTATTAATCTCCAGTAGTAGTACCAAGATTTGAAGAA 747
 Db 90 GCCTTATTCATGAGATGACTTCTTAAGTGCAGGCTCTACAGATATTTGTATATG 31
 Oy 748 GTGCAAAAACATTAAGTATACCTTAG 776
 Db 30 CTAAATTAATAATGACATTCATACAG 2

RESULT 8
 CDS05792/c 1100 bp DNA linear GSS 26-JUL-2000
 LOCUS CDS05792

DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
 007H05 of library A from Tetradon nigroviridis, genomic survey
 sequence.

ACCESSION AL324368
 VERSION AL324368.1 GI:9557250
 KEYWORDS GSS: genome survey sequence.

SOURCE Tetradon nigroviridis.
 ORGANISM Tetradon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 1100)
 Roest, Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fzames,C., Wincker,P., Brothier,P., Queller,F.,
 Saurin,W. and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetradon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE
 REFERENCE 20296633
 2 (bases 1 to 1100)
 Crollius,H.R., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Fzames,C.,
 Fischer,C., Bouneau,L., Billault,A., Queller,F., Saurin,W.,
 Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the

TITLE

OY 272 GTTCATATTGAGCACTTGTATTCACCTGACCTGACCTTACCGCCGATTCAGATTCACG 331
 Db 476 CAGGTCTGGAAAACCTGTTTCACACCTGACGCTAGACTACAGAGAGACGCGGCGATTGTGG 417
 OY 332 TTGCTTATGGCTCTCTTGACGGGTAAACACAAATCCCTTCGGT---TTTGAAGTCCAAAGA 388
 Db 416 CGAGGAACAGAGATGCTCTATCCGAGACCCAGGTGACCGGACGACTTTGTCTCCCAAGA 357
 OY 389 AAGGAATATGTGTGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 439
 Db 356 AAGGACAGACGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 297
 OY 440 GAGTCAAGTATTGCAATGAGCTAGACCAAAAGACATTGAAATCCATACCTTACGGCCAGCAT 499
 Db 296 GAGACCAATATCTACGCTGACCTGCTCCAAACACATCAAGGTGACGTTGGGGCAAGCCT 237
 OY 500 TTGGAGAAATGTGCAATGATATAAAATTTGATTTCTTACCATATCTGCTTGAATTTTATC 559
 Db 236 TTGGGAATTTCCGATGATTTTGGAGAAATCTATCAACAAATTTGCCAGCTGTAATTTCTATC 177
 OY 560 TTTCCTTTGAAATTTCAATCCAAAGATTACATACAGGAAAGCTATACATGCTTTTC 619
 Db 176 TTGCTTTGAGAACTCCATCCAAAGACTACATGAGGAGAAATCAACGGGCCCTTTTG 117
 OY 620 TTGGTGGCTCTGACCTGTGTTGTCGGACCATCTAGGAGAAACTATGAGAATTAAT-T 678
 Db 116 TTGGTGGCTCTGACCTGTGTTGTCGGACCATCTAGGAGAAACTATGAGAATTAAT 57
 OY 679 CCACGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 709
 Db 56 CTTTCGATTCGTTTATTCACGTTAACGACT 26

RESULT 10
 CNS01V8Y
 LOCUS
 DEFINITION
 Tetradon nigroviridis genome survey sequence puc-ori end of clone 198K9 of library G from Tetradon nigroviridis, genomic survey sequence
 accession AL166808.1 GI:7806865
 version 1
 keywords
 source
 organism
 SS: genome survey sequence.
 Tetradon nigroviridis.
 Tetradon nigroviridis.
 Ektariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetradon.
 1 (bases 1 to 844)
 Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 Title
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
 unpublished
 2 (bases 1 to 844)
 Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
 unpublished
 3 (bases 1 to 844)
 Genoscope.
 Direct Submission
 Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetradon>.
 Location/Qualifiers
 1..844
 source
 organism="Tetradon nigroviridis"

BASE COUNT	199 a	190 c	205 g	245 t	5 others
ORIGIN	/db_xref="taxon:99883" /clone="198K09" /clone_1lib="G" /note="Genoscope sequence ID : CONG198AF05SP1-end : pnc-ori"				
Query Match	17.6%; Score 161.4; DB 12; Length 844;				
Best Local Similarity	56.2%; Pred. No. 8.36-34;				
Matches 319; Conservative	2; Mismatches 244; Indels 3; Gaps 1;				
QY	25 ACTACTATTCGTGGTGGGTGGCCATTGGGGCAGACCTTTGACCTTACATCTTCGCCAA 84				
DB	280 ACCACGGGTGGTGGTGGGTGGCCATTGGGGCAGACCTTTGACCTTACATCTTCGCCAA 339				
QY	85 GCATATGTTCAATCCAGAGATGCCATCTCAACACGACCGTTGACTCTGTCACAAATCC 144				
DB	340 GCTTACTATGTTGATAGACGGCTGTTTCATACAGCGGACAGAACTATTTTCACCGAAGA 399				
QY	145 CATCATGTTGATGATCCATCAGAGACATCAAGTATGGATCTGACAAATTTACTCAGCAA 204				
DB	400 ATGGGAGTCCGTTATTCATCATGAGAAATTTCTCCGACCTGTCATCTGCGCAGAG 459				
QY	205 GCTAGGCGCCACCTCCAGAAATGGATTTGGATGAAATTTGGATACCACTCCACTGCC 264				
DB	460 GAGGCGCCGTTGATCCATCATGAGAGGCGTGGTTCACATGAGAAATCTCTGCTCCGAC 519				
QY	265 CAAGAAGTGCATGAGCAGCTGTTTAACTGACATGCTGACATTCACGCGCTGATGAT 324				
DB	520 AGGTTCTTGGAATGAGCGCTGTTTCAACCTGACTGATTAATGCTGATGATGCTGAT 579				
QY	325 ATCCAGTGGCTATGCTTCTTGAACGATGATGACAAATCCCTTCTGTTGAAAGTCCA 384				
DB	580 ATTTCATGCTCATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639				
QY	385 AGCAAAAGAAATTTGGTGTGCTGGTGTGAGTAATCTGGAACCTTGACATGCCAGATTC 444				
DB	640 AGCAAAAGAAATTTGGTGTGCTGGTGTGAGTAATCTGGAACCTTGACATGCCAGATTC 699				
QY	445 AAGTATTCAATGAGCTAAGCAAAAGCATTTGATTCATCTACCTTACGGGCAAGCATTTGA 504				
DB	700 AATATCTCAAGAGCTGCTCAAGCACAATGAACATCACTTATGAGACCTGCTTTGGG 759				
QY	505 GAATATGTCAATGATTAATAATTTGATTCCTATCATATCTGCTGTAAATTTATCTTTC 564				
DB	760 GGCTA---CAAGTGGGAAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816				
QY	565 TTGTGAATTCATCCACAGGATTTACA 592				
DB	817 TTGTGAATTCATCCACAGGATTTACA 844				
RESULT 11					
LOCUS	CNS0200S 887 bp DNA linear GSS 14-MAY-2000				
DEFINITION	Tetradon nlgroviridis genome survey sequence 17 end of clone 156120 of library g from Tetradon nlgroviridis, genomic survey sequence.				
ACCESSION	AL208693.1 GI:7867512				
VERSION	GSS: genome survey sequence.				
KEYWORDS	Tetradon nlgroviridis.				
SOURCE	Tetradon nlgroviridis.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodontidae; Tetraodon.				
REFERENCE	Roest-Crolius H., Jallion O., Dasilva C., Fizes M., Fisher C., Bouneau L., Billault A., Queller F., Saurin W., Bernot A. and Weissenbach J.				
AUTHORS	Characterization and repeat analysis of the compact genome of the				
TITLE					

Db 431 AATTCAATTCAGACAGACATCACTACTGAGAAACCTGCGATATGSCCTTCATCTGTG 490
 |||||
 QY 628 TCTGTACCTGTGTGTGGACCATCTAGGAGAAATATGATATATTCAGACAGAT 687
 |||||
 Db 491 GCTACACCTGTGTGTGGGCGCCCTTCCTTAACTATGAGCGCTTTTACCCCGAG 550
 |||||
 QY 688 TCAATTCATCTGTGAGATTTATCTCCAGTACGAGCAAGATATCTGAAGAA 747
 |||||
 Db 551 TCAATTCATCTGTGAGATTTATCTCCAGTACGAGCAAGATATCTGAAGAA 610
 |||||
 QY 748 GTGCAAAAACATTAAGTTATCTTACTTACTTACTG 788
 |||||
 Db 611 CTGTATTAAGATGATTAATATGATGCAACATTCACCTG 651
 |||||

RESULT 13
 BJO50300/c 666 bp mRNA linear EST 06-DEC-2001
 LOCUS BJO50300 NIBB Mochli normalized Xenopus neurula library Xenopus
 DEFINITION laevis cDNA clone X1023015.3', mRNA sequence.
 ACCESSION BJO50300
 VERSION BJO50300.1 GI:17392988
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 Xenopodidae; Xenopus.
 1 (bases 1 to 666)
 Kitayama, A., Terasaka, C., Mochli, M., Ueno, N., Shin-I, T. and Kohara, Y.
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadasi Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers
 1..666
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="X1023015"
 /clone_1ib="NIBB Mochli normalized Xenopus neurula library"
 /rissue_type="whole embryo"
 /dev_stage="stage 15"
 BASE COUNT 178 a 116 c 157 g 215 t
 ORIGIN

Query Match 15.6%; Score 151.6; DB 10; Length 666;
 Best Local Similarity 54.5%; Pred No. 46-31;
 Matches 348; Conservative 0; Mismatches 284; Indels 6; Gaps 2;
 QY 251 CAACCTACACTCCCAAAAGATGGCATGTGACACTGTGTTAACTGACTGACTTAC 310
 |||||
 Db 666 CAAAGCCACATCCCAACACAGCGTTATGACAACTCATCTGACACATGTTTAC 607
 |||||
 QY 311 GCCGTATTCAGATATCAAGTGCTTATGGCTTCTTACGAGTAAAGCAAAATCCCTCG 370
 |||||
 Db 606 GTGCTGATCTACAGATATTTTACTCTTATGCTGCTGCTAAAGGATGATGAAAAAGAA 547
 |||||
 QY 371 TGTTCAGTGGCCAAAGAGAAATGTGTGTGAGTGTGATGATGAACTGAGACCTG 430
 |||||
 Db 546 ACTTCATCTATGCGCCCAAGACAAACATGATGAGCTGGGTAGTCACTGATGAAACCAA 487
 |||||
 QY 431 AGATCCAGAGATGATTTACATGAGTAAAGCAAAAGATGAAATCCATFACCTAG 490
 |||||
 Db 486 ACTCCAGAGATTAATATTTATGAGAGATTAAGCCACACCTTGTGTATATCTATG 427
 |||||

QY 491 GGCAAGCATTTGGAAATATGTCAATGATTAATAATTGATTCCTACCATCTGCTGTA 550
 |||||
 Db 426 GCGTACAGCATTTACCCCTTCCTGAGACCAAAACA---CATGAAATTTCTTCCCATATA 370
 |||||
 QY 551 AATTTATCTTCCCTTGGAAATTCATCCCAAGATTAATACAGGAAACCTAT--- 607
 |||||
 Db 369 AATTTATTTAGCTTTTGGAAATTCATTCACAGACTCATCATCATGAACTGTGGC 310
 |||||
 QY 608 ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 667
 |||||
 Db 309 ATAAAGCCCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
 |||||
 QY 668 AGAATTAATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 727
 |||||
 Db 249 AGCGTTATTTTACCCCGAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 190
 |||||
 QY 728 TACCAAGTATCTGAGAGAGATGACCAAAACATTAAGTTATACCTTACTTACTTACT 787
 |||||
 Db 189 TTGCTGATTTATCTAAACAGACCTTGATTAAGATATATAATACATGCAACATTTCACT 130
 |||||
 QY 788 GGAGAGAGATTTTACCTGTAATCTCCAGATTTTGGAGATGATGATGATGATGATGATG 847
 |||||
 Db 129 GGAGCAAAACATCTAAACACAGTACACCTGAGAGTTGGAACCCCATTTCTGCAAGTAT 70
 |||||
 QY 848 GCGATCATGTGAAAGGACATCAAGATTAATTAAGTCTGT 885
 |||||
 Db 69 GCAAGCTATTAATAAGACGACACCGCTTATAGACTTTT 32
 |||||

RESULT 14
 CENS034XF/c 711 bp DNA linear GSS 15-MAY-2000
 LOCUS CENS034XF tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 DEFINITION 212J18 of library 6 from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL228012
 VERSION AL228012.1 GI:7886987
 KEYWORDS GSS; genome survey sequence.
 ORGANISM Tetraodon nigroviridis.
 Tetraodon nigroviridis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 711)
 Roest-Crolius, H., Jallion, O., Dasilva, C., Fizesmes, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetler, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished
 2 (bases 1 to 711)
 Roest-Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizesmes, C., Wincker, P., Brothier, P., Quetler, F.,
 Saurin, W. and Weissenbach, J.
 Human gene number estimate provided by genome wide analysis using
 tetraodon nigroviridis DNA sequence
 Unpublished
 3 (bases 1 to 711)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 Location/Qualifiers
 1..711
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="212J18"
 /clone_1ib="G"
 /note="Genoscope sequence ID : C0AG212DE09SP1-end :

FEATURES
 source
 1..711
 /db_xref="taxon:99883"
 /clone="212J18"
 /clone_1ib="G"
 /note="Genoscope sequence ID : C0AG212DE09SP1-end :

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 18.6694 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-4_COPY_1454_2368

Perfect score: 915
Sequence: 1 accaaactgattatttaa.....agaatgcttcgtaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	13.9	1814	2	US-08-483-151-1
2	127	13.9	1814	5	PCR-US96-06427-1
3	113.4	12.4	1654	1	US-07-914-281-13
4	113.4	12.4	1654	1	US-08-393-246-13
5	113.4	12.4	1654	1	US-08-325-058A-13
6	113.4	12.4	1654	2	US-08-696-731-13
7	113.4	12.4	1654	4	US-09-042-531-13
8	112.2	12.3	1086	1	US-07-914-281-12
9	112.2	12.3	1086	1	US-08-393-246-12
10	112.2	12.3	1086	1	US-08-525-058A-12
11	112.2	12.3	1086	2	US-08-696-731-12
12	112.2	12.3	1086	4	US-09-042-531-12
13	112.2	12.3	2042	3	US-09-063-237-2
14	112.2	12.3	2043	1	US-07-914-281-1
15	112.2	12.3	2043	1	US-08-393-246-1
16	112.2	12.3	2043	1	US-08-273-411-4
17	112.2	12.3	2043	1	US-08-525-058A-1
18	112.2	12.3	2043	2	US-08-696-731-1
19	112.2	12.3	2043	4	US-09-042-531-1
20	112.2	12.3	2043	5	PCR-US91-00899-1
21	111.8	12.2	1316	1	US-07-914-281-10
22	111.8	12.2	1316	1	US-08-393-246-10
23	111.8	12.2	1316	1	US-08-525-058A-10
24	111.8	12.2	1316	2	US-08-696-731-10
25	110.8	12.2	1316	4	US-09-042-531-10
26	100.8	11.0	1215	5	PCR-US91-00899-5
27	100.8	11.0	1215	5	PCR-US91-00899-5

28	100.8	11.0	1488	1	US-07-914-281-9	Sequence 9, Appl
29	100.8	11.0	1488	1	US-08-393-246-9	Sequence 9, Appl
30	100.8	11.0	1488	1	US-08-525-058A-9	Sequence 9, Appl
31	100.8	11.0	1488	2	US-08-696-731-9	Sequence 9, Appl
32	100.8	11.0	1488	4	US-09-042-531-9	Sequence 9, Appl
33	100.8	11.0	2175	4	US-08-482-073-9	Sequence 9, Appl
34	100.8	11.0	2861	4	US-08-482-073-10	Sequence 9, Appl
35	100.8	11.0	3647	1	US-07-914-281-7	Sequence 10, Appl
36	100.8	11.0	3647	1	US-08-393-246-7	Sequence 7, Appl
37	100.8	11.0	3647	1	US-08-525-058A-7	Sequence 7, Appl
38	100.8	11.0	3647	2	US-08-696-731-7	Sequence 7, Appl
39	100.8	11.0	3647	5	PCR-US91-00899-4	Sequence 7, Appl
40	100.8	11.0	3647	5	PCR-US91-00899-4	Sequence 7, Appl
41	97.6	10.7	2134	2	US-08-483-151-3	Sequence 4, Appl
42	97.6	10.7	2134	5	PCR-US96-06427-3	Sequence 3, Appl
43	37.2	4.1	1728	1	US-08-403-866-12	Sequence 12, Appl
44	37.2	4.1	2231	1	US-08-403-866-11	Sequence 14, Appl
45	37.2	4.1	12720	1	US-08-403-866-11	Sequence 11, Appl

ALIGNMENTS

Result 1
US-08-483-151-1
; Sequence 1, Application US/08483151
; Patent No. 5858752
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Holgersson, Jan
; TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,151
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/278001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-483-151-1
Query Match 13.9%; Score 127; DB 2; Length 1814;
Best Local Similarity 52.3%; Pred. No. 9.7e-29;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
QY 104 GATGCATCTCAACAGCGCTTCACTGATACCAACATCCATGCGTTCGATGCATC 163
DB 548 GCTGCCGTGTGAGTGTCAACCGAGGCTCTAGACCGAGTGTGCTGTGCTTTCACCC 607

Tue Oct 8 10:18:29 2002

us-09-744-748-4 copy 1454 2368.rni

Page 2

OY	164	ACGAGACATCATAGTTGGGATCTGACAAATTTTACCTCAGCAGCTTAGGCCACCCCTTCAG	223
Db	608	ACCGGTGAGCTGCAAAACCCGGCAATCTCTCTACCCCTGGACAGAGGCGACACGGACGC	667
OY	224	AATGGATTTTGATGAAATTTTGGAAATCAACCACTCACACTCCCCAAAAGAGTGGCATTTAGC	283
Db	668	CTTGGGTGTGGGCTCCATGGAATGAGCCCAAGTAATACCATGATGCTCCATGCGCTTCCGGG	727
OY	284	ACTGTTTAAACCTGACCTGTGACTTACCGCGCGTATTCAGATATCCAGTGCCTTATGACT	343
Db	728	GCACTCTTCAACTGGGTGCTGACGTATTCAGTACCGCGTGAATTCAGATATCTTTGTATCCCTACGGTTC	787
OY	344	TCCTTGACGGTAGACACAAATCCCTTGTGTGTGGAAGTGGCAAGCAAGAAGAAATGGTGT	403
Db	788	GCTTTGAGAGCTCTCTCTGCGGGCC---CACATGCCCATACCGGCGCCAAAAGCAGATGGCGTG	844
OY	404	GCTGGGTGTGAGTAACCTGGAACCTCGACACATGCCAGATCAAGTAAATTACATGAGCTAA	463
Db	845	CGTGGTGATACGAATTTTCCAGAGCGGAGAGAGGTGCAAAGCTGATACGGCAGCTGG	904
OY	464	GCAAAACATTTGAATTCATACCTACCGGCGCAAGCATTTGGAGAAATATGTACATGATAAA	523
Db	905	CCCCCATCTGCAAGTGTGATCTTGTGGTCCGCGCCAGGCGGAGCGGCCCTATGCGGTATTT	964
OY	524	ATTGATTTCTTACCATATCTGCTGTGTAATTTTATCTTCTTGTAAATTCATATCCACA	583
Db	965	GTCGTGCTGCCACTTTTGGCCCGGTACCGCTTCTACCTGGCCTTTGAGATCACAGCATC	1024
OY	584	AGATTTACATCAGCGGAAAGCTAT--ACAAATCTTTCTGGCTGGCTGTGTACTGTTG	640
Db	1025	GGGCATCATCATGAGAAATTTGTGGCGCAATGCCCTGGCGGCTGTGCTGATCCCGGG	1084
OY	641	TTCTGGGACCATCTAGGAGAAACTATGAGAAATTAATTTTCACAGACATTCATTCATTCATG	700
Db	1085	CGGTGGAGCCCTCCCGGGCCACCTACAGAGGCTTTGTGGCACAGATGCTTTGTGACAG	1144
OY	701	TGGAAGATTATTACTCTCCAGTAGACTAGC	731
Db	1145	TGACAGACTTACACTCTGCCCCGTAAGACTGC	1175

TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1814 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US96-06427-1

Query Match 13.9% Score 127 DB 5 Length 1814;
Best Local Similarity 52.3% Pred. No. 9,7e-29;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

OY 104 GATGCCATCTCAACAGCGAGCGTTCATCGTATACACAAATCCCAAGAGTTCGATGATC 163
Db 548 GCTGCCCTCGAGTGGCTAACCAGACCTCGTAGCCAGTGTGCTGCTTCACAC 607
OY 164 ACCGAGCAATCACTGGATGGATCTGCAGATTATTAAGTACAGAGTACGCCACCTTCACAG 223
Db 608 ACCGTGAGCTGCACAAACCGGCAATCTCTCTACCCCTGAGACAGAGGACACAGGACAGC 667
OY 224 AATGATTTGGATGAATTTGGAAATCACCACACTACCTCCCAAAAGAGTCAATGAGC 283
Db 668 CTGGGGCTGGCGCTCCATGATATCGCCAGATATACCCATGCTTCATCGCTTCGCGG 727
OY 284 ACTGTTTAACCTGACTGACTTACCTACCGCGCGTGAATTCAGATATCCAAATGCTATAGCT 343
Db 728 GCATCTTCAACTGGGTGGTGGATCTGAGCTATGGCGCTGATTCAGATATCTTGTACCTAGGTC 787
OY 344 TCTTGACGGTAAACCAAAATCCCTTCGTGTTTGAAGTGCACAGCAAAAGAAATATGCTGT 403
Db 788 GCTTGGAGCCTCTCTCTGTGGGCC---CACATCCCACTACCGGCCAAACAGAGATGGCTG 844
OY 404 GCTGGGTGTGATGTAAGTCTGGAACCCCTGAGCATGCCAGATCAAGTATTAACAATGAGCTAA 463
Db 845 CCTGGGTGATACACAATTTCCAGAGCGGACAGACAGCTGCAAAAGCTGATCCGGCAGCTGG 904
OY 464 GCAAAACATTTGAAATCCATACCTACGCGGCAAGATTTGGAGAAATATGCAATGATGATAA 523
Db 905 CCCCTCATCTGCAGGTGATGTGGTGGTCCGCGACAGGCGAGCCCTATGGGCTAAT 964
OY 524 ATTGATTCCTACCATATCTGCTTGAATTTATCTCTTTGAAATTTCAATCCACAA 583
Db 965 GTCTGCGCCCACTTTGGCGCGGTACCCGCTTCACTGCGCTTTGAGAACTCAGACATAC 1024
OY 584 AGAATTAATCAACGAGAAAGCAT---ACAATGCTTTTGGCGTGGCTGCTGATACCTGGT 640
Db 1025 GGGAGTCACTACCTGAGAACTTCTGGCGCAATGCCCTGGCGGTGCTGATACCCGTGG 1084
OY 641 TTCTGGGACCATCTAGGAGAAACTATGAAATATATTTCCAGACAGATCTATTCATTCATG 700
Db 1085 CGCTGGACCTCTCTGGGCGACCTTACGAGGCTTTTGTGCAACAGATGCTTTGTACAG 1144
OY 701 TGGAGATTATACCTCTCCAGTGAAGTAC 721
Db 1145 TGGACGACTTCACTCTGCCCGTGAAGTGGC 1175

RESULT 3
US-07-914-281-13
Sequence 13, Application US/07914281
Patent No. 5324663
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C. Jefferson Davis Highway, Fourth Floor
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 1992/07/20
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-914-281-13

Query Match 12.4%; Score 113.4; DB 1; Length 1654;
Best local Similarity 51.2%; Pred. No. 1.2e-24;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
DB 31 ATTCTGTGTGGTGGCCATTGGGCGAGACCTTTGACCTTACCTTCGCAAGCAATG 90
DB 318 ATCTCTGTGGAGCGGCTTTTACCAACCCATAGCTGCGCGCTGCTAGAGATG 377
DB 91 TT---CAACATCCAGAGTCCATTCACAGGACCGCTTACCTTACCTTACCAATCCAT 147
DB 378 GTCTGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
DB 148 GCAATTTCTGATCCATCACCAGACATCAGTTGGATCTGA---CAATTTACCTCAGCAA 204
DB 438 GCGGTATCTGTGACACCGAGAGGTGATGATGATGATGATGATGATGATGATGATGATG 457
DB 205 GCTAGGCGACCTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 264
DB 498 CCGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 557
DB 265 CAAAGAGTGGATGAGCATGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 324
DB 558 CACTGTAAGACCTTGAAGCATGCTTCAATCTCAACATGCTTCAACGACGCTGCTGCTGCT 617
DB 325 ATCCAGTGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375
DB 618 ATCTTACGCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
DB 376 GAAGTGGCAAGCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 435
DB 678 AACCTTCCGCGCAAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
DB 436 GCGAGGTCAAGTATTAACATGAGCTAAGCAAAAGATTAATTCATCTACTACCTGAGGCA 495
DB 738 GCGAGGTGCGCTACTACGACGACCTGCGAGGCCCATCTCAAGGTGGCGCTGCTGCTGCTGCT 794
DB 496 GCAATTTGAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
DB 795 CGCTCCCAAGAGCCCTGCGGCCGAGGGAACATGATGAGAGCGCTGCTGCTGCTGCTGCTGCT 854

DB 556 TATCTTCTTTGAAAATTAATCCACAGAGATTACATCCGAAAAGCTAT---ACAAT 612
DB 855 TATCTGCTTTCAGAACTCTTCCACCCGACTACATCCAGAGAGCTGTGAGAGCA 914
DB 613 GCTTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
DB 915 GCCCTGAGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 974
DB 673 TATATTCACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 732
DB 975 TTCTGCGACCGCGCGCTTACATCCAGAGAGCTTCCAGAGGCCCAAGAGAGCTGCGCC 1034
DB 733 AAGTATTCGAGAGATTCGCAAAAACATTAATTAATTAATTAATTAATTAATTAATTAATTA 792
DB 1035 CGGTACTCTGAGAGCTGAGCAAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094
DB 793 AAG 795
DB 1095 GAG 1097

RESULT 4
US-08-393-246-13
Sequence 13, Application US/08393246
Patent No. 5535900
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS.
TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-393-246-13
Query Match 12.4%; Score 113.4; DB 1; Length 1654;
Best local Similarity 51.2%; Pred. No. 1.2e-24;

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

```

OY 31 ATTCTGCTGGTGGTGGCCATTGTTGGGAGACCTTTGACCTTACATCCGCCAAGCAANG 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ATCTGCTGGTGGTGGCCATTGTTGGGAGACCTTTGACCTTACATCCGCCAAGCAANG 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 91 TT---CAACATCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTGCTGGGACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 148 GCAGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GGGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 205 GCTAGGCCACCTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CCGAGCGCGGAGGGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 265 CAAAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CAGCTGAAAGCCATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 325 ATCCAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 ATCTTCAGCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 376 GAACTGCCAAGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AACCTCTGGCCAAAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 436 GCCAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCCAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 496 GCATTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 GCCTGCCAAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 556 TATCTTCTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 TATCTTCTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 613 GCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 915 GCTCTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 673 TATATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 975 TTCTGCGCACCCGAGCGCTTATCAGCTGAGAGACTTCAGAGCCCAAGAGACCTGGGC 1034
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 733 AAGTATCTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 CGGTACTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 793 AAG 795
    |||
Db 1095 GAG 1097
    |||

```

```

RESULT 5
US-08-525-058A-13
; Sequence 13, Application US/08525058A
; Patent No. 5770420
;
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATHER & NEUSTADT, P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor

```

```

; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)486-2347
; TELEFAX: (703)486-2347
; TELETYPE: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-525-058A-13

```

Query Match 12.4%; Score 113.4; DB 1; Length 1654;

Best Local Similarity 51.2%; Pred. No. 1.2e-24;

```

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

OY 31 ATTCTGCTGGTGGTGGCCATTGTTGGGAGACCTTTGACCTTACATCCGCCAAGCAANG 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ATCTGCTGGTGGTGGCCATTGTTGGGAGACCTTTGACCTTACATCCGCCAAGCAANG 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 91 TT---CAACATCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTGCTGGGACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 148 GCAGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GGGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 205 GCTAGGCCACCTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CCGAGCGCGGAGGGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 265 CAAAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CAGCTGAAAGCCATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 325 ATCCAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 ATCTTCAGCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 376 GAACTGCCAAGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AACCTCTGGCCAAAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 436 GCCAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCCAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 496 GCATTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CGCTGCCAAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 556 TATCTTCTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 TATCTTCTTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Tue Oct 8 10:18:29 2002

us-09-744-748-4_copy_1454_2368.rni

Page 6

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEPHONE: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-13

Query Match 12.4%; Score 113.4; DB 4; Length 1654;
Best Local Similarity 51.2%; Pred. No. 1,2e-24;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
QY 31 ATTCTGTGGTGGGTCATTTGGGAGACGCTTTGACCTTACATCTTCAAGCAAG 90
DB 318 ATCTCTGTGGAGCGGCTTTTAAACCAACCACTGCTGCCCCCTGCTCAGAG 377
QY 91 TT---CAACATCCAGATGCCATCTCAACAGACGCTTACCTGTACAAATCCAT 147
DB 378 GTGCTGCGACGGCTGACTGCAACATCTCCGACGCAAGGTGTATCCACAGGAC 437
QY 148 GCATTTCTGATCATCACCAGACATCAGTTGGATCTGA---CAATTTACCTCAGCA 204
DB 438 GCGGTATGTCACACCGACGAGAGGTCTATCAACCCGAGTCCGACGCTCC 497
QY 205 GCTAGGCAACCTTCAGAAATGATTTGATTTGATTTGATTTGATTTGATTTGAT 264
DB 498 CCGAGGCGGACGAGGCGAGGATCTGTTACAGATGAGATGAGATGAGATGAGAT 557
QY 265 CAAGAAGTGGATGAGACTTGTTTACCTGACTGACTGACTGACTGACTGACTGACT 324
DB 558 CACCTGAAGGATGAGAGGATTAATCTCAACATGCTTCAACGCGACGACTCGAC 617
QY 325 ATCCAGTGGCTTATGCGT-----TCTTGAAGGGAAGCAACATCCCTTCGT 375
DB 618 ATTTTACGCGCTTACGCGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGG 677
QY 376 GAAGTCCCAAGCAAGAAATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
DB 678 AACCTCTGCGGCAAGACGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
QY 436 GCCAGAGTCAATATTATGATGAGTAAAGCAAGATTTGAATCCATTACTAGGGCAA 495
DB 738 GCCAGGAGTCTGCTACTACAGAGCTGAGGCGGCTGCAAGGTGAGAGCTGTACG 794

QY 496 GCAATTTGAGAAATATGATGATTAATAATTTGATTTCCACCATATCTGTAATAT 555
DB 795 CGTCCACACAGCCCTGCCCCAGGAGACCATGATGAGAGCTGTCCCGTCAAGATTC 854
QY 556 TATCTTCTTTGAAATTCATCCACAAAGATTTATCTACGAGGAAAGCTAT---ACA 612
DB 855 TATCTGCGCTTCAGAACCTCTGACCCCCGACTCATACACGCAAGACCTGTGAGAAC 914
QY 613 GCTTTTGTGCTGCTCTGTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 672
DB 915 GCGCTGAGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
QY 673 TATATTCACGACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 732
DB 975 TTCTGCGACCGGACGCTTCATTCATTCATTCATTCATTCATTCATTCATTC 1034
QY 733 AAGTATCTGAAGAGTGCACAAACAAATTAATTAATTAATTAATTAATTAATTA 792
DB 1035 CGGTACTGCGAGGAGCTGAGACAGCACGCGGCTGCTGCTGCTGCTGCTGCTG 1094
QY 793 AAG 795
DB 1095 GAG 1097

RESULT 8
US-07-914-281-12
Sequence 12, Application US/07914281
Patent No. 532463
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OHLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-914-281-12
Query Match 12.3%; Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2,2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;


```

OY 106 TGGCATCTCACAAGGACCGCTTCACTGTCACAAATCCGACGCTTGATCATCAC 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 271 TGGCACAATCATGCGCCGCGCAAGGTGTACCCACAGGACAGACGGTCACTCGACACAC 330
OY 166 CGAGCATCATGTTGGG---TCTGACAAATTTACCTCAGACAGCTGGCCACCCCTTCAG 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 331 TGGGATATCATGTCACACCCCTTAAGTAGAGCTCCACCTTCCCGGAGCCGCGAGGGGAG 390
OY 223 AATGATTTGGATGATTTGGATTCACCAACTCACACTCCCAAAAAGAGTGGCATTTGAG 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 CGGTGATCTGTTCACTTGGAGCCACCCCTTAAGTGGAGAGCTGGAGAGCCCTGGAG 450
OY 283 CACTGTTTAACTGACTGCTGACTACCGCCGCTGATTCAGATATTCAGAGCTTATGCG 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 AGATATCTCAATCTCACCATGTCCTACCCGAGGAGCTCCGACATCTTCACGCGCTACGCG 510
OY 343 TTCTTACGGTAAAGCAAAATCCCTTGT-----GTTGAATGCCAAGCAAAAGAG 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 511 TGGCTGAGCCGCTGTGCGGCGCAGCCCTGCCACCCGCTCAACTCTGCGCCAGAGC 570
OY 394 AATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 571 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 630
OY 454 AATGACTTAAGCAAAAGCATTTGAATTCATACCTACGCGGCAAGCATTTGGAGATATGTC 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 631 CAGAGCTGCTGAGGCTCATCTCAAGGTGAGCTGTCAGGAGCTCCACAAAGCCCTGCGCC 690
OY 514 AATGATTAATAATTTGATTTCTTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 573
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 691 AAGGGGACCA---TGATGAGAGAGCTGTCCCGTACAAATTTCTACTGCGCTTCGAGAAC 747
OY 574 TCAATCTCACAAGGATTAATCATACGCAAAAGCTAT---ACAATGCTTTTCTGCTGCTG 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 748 TCTCTTGCACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807
OY 631 GTAACCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 808 GTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
OY 691 TTCATTCATGTTGAGATATTAATCTCCAGAGGAGCTGCAAAAGTATCTGAAGAGATC 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 868 TTCATTCATGTTGAGATATTAATCTCCAGAGGAGCTGCAAAAGTATCTGAAGAGATC 927
OY 751 GACAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 928 GACAAAGACACGCGCTACCTGAGCTACTTTCGCTGCGGAG 972

```

```

RESULT 9
US-08-393-246-12
: Sequence 12, Application US/08393246
: Patent No. 5595900
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B.
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
: TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
: TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUBAUER,
: ADDRESSER: P.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/393,246
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/220,433
: FILING DATE: 30-MAR-1994
: APPLICATION NUMBER: US 07/914,281
: FILING DATE: 20-JUL-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Lavalleye, Jean-Paul M. P.
: REGISTRATION NUMBER: 31,451
: REFERENCE/DOCKET NUMBER: 2363-060-55
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)521-4500
: TELEFAX: (703)486-2347
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1086 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: US-08-393-246-12

```

Query Match 12.3%; Score 112.2; DB 1; Length 1086;

Best Local Similarity 51.6%; Pred. No. 2.2e-24;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

OY 106 TGGCATCTCACAAGGACCGCTTCACTGTCACAAATCCGACGCTTGATCATCAC 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 271 TGGCACAATCATGCGCCGCGCAAGGTGTACCCACAGGACAGACGGTCACTCGACACAC 330
OY 166 CGAGCATCATGTTGGG---TCTGACAAATTTACCTCAGACAGCTGGCCACCCCTTCAG 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 331 TGGGATATCATGTCACACCCCTTAAGTAGAGCTCCACCTTCCCGGAGCCGCGAGGGGAG 390
OY 223 AATGATTTGGATGATTTGGATTCACCAACTCACACTCCCAAAAAGAGTGGCATTTGAG 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 CGGTGATCTGTTCACTTGGAGCCACCCCTTAAGTGGAGAGCTGGAGAGCCCTGGAG 450
OY 283 CACTGTTTAACTGACTGCTGACTACCGCCGCTGATTCAGATATTCAGAGCTTATGCG 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 AGATATCTCAATCTCACCATGTCCTACCCGAGGAGCTCCGACATCTTCACGCGCTACGCG 510
OY 343 TTCTTACGGTAAAGCAAAATCCCTTGT-----GTTGAATGCCAAGCAAAAGAG 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 511 TGGCTGAGCCGCTGTGCGGCGCAGCCCTGCCACCCGCTCAACTCTGCGCCAGAGC 570
OY 394 AATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 571 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 630
OY 454 AATGACTTAAGCAAAAGCATTTGAATTCATACCTACGCGGCAAGCATTTGGAGATATGTC 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 631 CAGAGCTGAGGCTCATCTCAAGGTGAGCTGTCAGGAGCTCCACAAAGCCCTGCGCC 690
OY 514 AATGATTAATAATTTGATTTCTTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 573
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 691 AAGGGGACCA---TGATGAGAGAGCTGTCCCGTACAAATTTCTACTGCGCTTCGAGAAC 747
OY 574 TCAATCTCACAAGGATTAATCATACGCAAAAGCTAT---ACAATGCTTTTCTGCTGCTGCT 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 748 TCTCTTGCACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807
OY 631 GTAACCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 808 GTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
OY 691 TTCATTCATGTTGAGATATTAATCTCCAGAGGAGCTGCAAAAGTATCTGAAGAGATC 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 868 TTCATCCAGCTGGACACTTCACAGAGCCCAAGACCTGGCCGCTACCTGACAGAGCTG 927
QY 751 GACAAAACATTAAGTTATCCCTTACTTACTTACTTACTGAGAG 795
Db 928 GACAGAGACCAACGCCCGCTACCTGACTTCTTCCGTGGCGGAG 972

RESULT 10
US-08-525-058A-12
Sequence 12, Application US/08525058A
Patent No. 5770420
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525, 058A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-525-058A-12

Query Match 12.3%; Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;
QY 106 TGGCATCTCAACGACGCTTCCTGACCAACAAATCCCATGAGCTTGTGATCCATGAC 165
Db 271 TGGCATCTCAACGACGCTTCCTGACCAACGACGACGACGACGACGACGACGACGAC 330
QY 166 CGACACATCATGTGGGCA---TTCGACAATTTACCTCAGCAGTACGACGACGACGACGAC 222
Db 331 TGGCATCTCAACGACGCTTCCTGACCAACGACGACGACGACGACGACGACGACGACGAC 390
QY 223 AATGATTTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 282
Db 391 CGCTGATCTGCTTCACTTGGAGCCACCCCTTAAGTGGCAGACCTGGAAGCCCTGGAC 450
QY 283 CACTTGTTAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 342
Db 451 AGACTTCAATCTCAACATGCTCAACGACGACGACGACGACGACGACGACGACGACGACGAC 510
QY 343 TTCTGACGCTAAGCAACAAATCCCTTCGT-----GTTTGAAGTCCCAAGAAAGAG 393

Db 511 TGGCTGAGACCGCTGGTCCCGGACAGCTCCACCCACCCCTCAACTCTTCGGCCCAAGACC 570
QY 394 AATTTGCTGTGCTGGGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
Db 571 GAGCTGTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 630
QY 454 AATGAGCTAAGCAAAAGCATTTGAATCCATACCTACGAGGAGCAATTTGAGATATGTC 513
Db 631 CAGAGCTGACAGCTCATCTCAGGTGACGCTGACGAGCTCCACAAAGCCCTGACC 690
QY 514 AATGATTTAAATTTGATTTCTACCATATGCTGTTGATTTATCTTCTTTGAAAT 573
Db 691 AAGGGGACCA--TGATGAGACAGCTGTCCCGGATCAAGTTCTACCTGAGCTTCGAGAAC 747
QY 574 TCAATCCACAGATTTACATCAGGAAAGCTAT---ACAATGCTTTTGTGGCTGCTCT 630
Db 748 TCTTGACACCCGACTACATCATCAGAGAGCTGTGAGAGAACGCCCTGGAGGCTGTGGCC 807
QY 631 GTACCGTTGTTCTGGGACCACTAGAGAAACTATGAGATATATTCACAGATATCA 690
Db 808 GTGCCCGTGTCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 867
QY 691 TTCATTCATGTGGAAGATTTACTTCTCCAGTACGCTACCAAGATTTCTGAAGAGTTC 750
Db 868 TTCATTCATGTGGAAGATTTACTTCCAGAGCCCTGACGCTGACGCTGACGCTGACGCTGAC 927
QY 751 GACAAAACATTAAGTTATCCCTTACTTACTTACTTACTTACTTACTTACTTACTTACT 795
Db 928 GACAGAGACCAACGCCCGCTACCTGACTTCTTCTGCTGGCGGAG 972

RESULT 11
US-08-696-731-12
Sequence 12, Application US/08696731
Patent No. 5955347
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696, 731
FILING DATE: 14-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/393, 246
FILING DATE:
APPLICATION NUMBER: US 08/220, 433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914, 281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-696-731-12

Query Match 12.3%; Score 112.2; DB 2; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 2.2e-24;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

106 TGCATCTCAACAGGACGCTTCACTACACAAATCCATGCTGATCATCAC 165
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 271 TGGCATATCATGCCAGCGCAAGGTGTACCCAGCAGCAGCGTCACTGACAC 330
 166 CGAGACATCAGTTGGGA---TCTGACAAATTTACCTCAGCAAGCTAGGCCACCTTCAG 222
 331 TGGGATATCATGTCCAAACCTAAGTCACGCTCCACCTTCCCGAGGCGCAGGGGACAG 390
 223 AATGATTTGGATGATTTGGATACCACTACACACCTCCCAAAAGAGTGCATTGAG 282
 391 CGCTGATCTGTCTCACTTGAAGCCACCCCTTAAGCTGCAGCAGCTGGAAGCCCTGAGC 450
 283 CACTTGTTAACCTGACTGACTGACTACCGCCGCTGATTCAGATTCCTCAAGTGCCTTAGC 342
 451 AGATATTCTCATCTCATCATCTGCTTACCCAGCAGCAGCTCCGACATCTTCACGCCCTTAGC 510
 343 TTCTTGACGCTAAGCACAATTCCTTCGT-----GTTGAAGTGCACCAAGCAAGAG 393
 511 TGGCTGAGACCGCTGGTCCGACGACCTGACCCACCCGCTCAACCTCTGAGCCAGAC 570
 394 AATGATGCTGCTGGTGTGAGTACGTAACCAACCTGAGACATGCCAGAGCAATATAC 453
 571 GAGCTGTGCTGGCTGGGCTGTCCAACTGCAAGCCGACTCAGCAGGCTGCTACTAC 630
 454 AATGACCTAAGCAAAAGCAATTTGAAATTCATACCTAGGCGCAACATTTGGAATATGTC 513
 631 CAGAGCTGCAAGCTCATCTCAAGGTGAGCTGTACGAGCGCTCCCAAGACCCCTGCCC 650
 514 AATGATAAAAATTTGATTCCTACCATATCTGCTTGAATTTATCTTCCCTTTGAAAT 573
 691 AAGGGGACCA---TGATGAGAGCGCTGCTCCGGTACCAAGTTCTACCTGCGCTTCGAGAAC 747
 574 TCAATCCACAAGATATACATACAGGAAAAGCTAT---ACAATCTTTCTGCGTGCCTC 630
 748 TCCCTGACCCCGACTACATACCGAGAAAGCTGTGAGGAACCCCTGAGGCGCTGGGCC 807
 631 GTACCTGTTGTTCTGGGACATCTAGAGAAAATATATATATTCAGAGCATTTCA 690
 808 GTCCCGGTGTGTGTGGGCCCCCAGCAAGCAACTAGAGAGGTCTCTGCCACCCGACGCC 867
 691 TTCATTCATGTGGAGATTTATACTCTCCAGTGAAGTATCTGAAGAGTGC 750
 868 TTTCATCAGTGGAGCACTTCCAGAGCCCAAGGACCTGCGCGTACTCTGAGAGCTG 927
 751 GACAAAACAT 795
 928 GACAAAGCACCGCCGCTACTGAGCTACTTTCGCTGGCGGGAG 972

RESULT 12

US-09-042-531-12
 Sequence 12, Application US/09042531
 Patent No. 6268193

GENERAL INFORMATION:
 APPLICANT: LOWE, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS.
 TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLION, SPIVAK, MCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/042,531
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:
 APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994
 APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-09-042-531-12

Query Match 12.3%; Score 112.2; DB 4; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 2.2e-24;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

106 TGCATCTCAACAGGACGCTTCACTGATACAAATCCATGCTTGCATCATCAC 165
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 271 TGGCATATCATGCCAGCGCAAGGTGTACCCAGCAGCAGCGTCACTGACACAC 330
 166 CGAGACATCAGTTGGGA---TCTGACAAATTTACCTCAGCAAGCTAGGCCACCCCTTCAG 222
 331 TGGGATATCATGTCCAAACCTAAGTCACGCTCCACCTTCCCGAGGCGCAGGGGAC 390
 223 AATGATTTGGATGATTTGGATACCACTACACACTGACCTCCCAAAAGAGTGCATTGAG 282
 391 CGCTGATCTGTTCAACTGAGGACACCCCTTAAGTGCAGACACTGGAAGCCCTGAGAC 450
 283 CACTTGTTAACCTGACTGACTGACTACCGCCGCTGATTCAGATTCCTCAAGTGCCTTAGC 342
 451 AGATATTCTCATCTCATCATCTGCTTACCCAGCAGCAGCTCCGACATCTTCACGCCCTTAGC 510
 343 TTCTTGACGCTAAGCACAATTCCTTCGT-----GTTGAAGTGCACCAAGCAAGAG 393
 511 TGGCTGAGACCGCTGGTCCGACGCTGACCCACCCGCTCAACCTCTGCGCAAGAC 570
 394 AATGATGCTGCTGGTGTGAGTAACTGGAACCTGAGCATGCAAGATCAAGTATATAC 453
 571 GAGCTGTGCTGGGCGGTGTCCAACTGGAAGCCGAGACTCAGCAGGCTGCTACTAC 630
 454 AATGACCTAAGCAAAAGCAATTTGAAATTCATACCTAGGCGGCAAGCATTTGGAATATGTC 513
 631 CAGAGCTGCAAGCTCATCTCAAGGTGAGCTGTACGAGCGCTCCCAAGACCCCTGCCC 690

QY 514 AATGATAAAATTTGATTCCTACCATATCTGTTGTAATTTATCTTCCCTTGAANA 573
DB 691 AAGGGAGACA--TGATGAGAGAGCTGTCCGGTACAAAGTTCTACTGGCTTCGAGAAC 747
QY 574 TCAATCCAAAGATTTACATCAGGAGAAAGCTAT---ACAATGCTTTCTGCTGCTCT 630
DB 748 TCCCTGACACCCGACTCATCTACGAGAACCTGTGAGAGAACCCCTGAGAGCTGAGGC 807
QY 631 GTACCTGTCTTCGAGACCATCTAGGAAACTATGAGATTTATTCAGCAGATTC 690
DB 808 GTCCCGCTGTGCTGGAGCCAGAGACACTAGAGAGTTCTGCTGAGCCAGCCGAGGC 867
QY 691 TTCAATCATGTGAGAGATTTATCACTCTCCAGTAGAGCTAGCAAAAGTATCTGAAGAGTC 750
DB 868 TTCAATCCAGCTGAGAGACTTCCAGAGCCCAAGAGACCTGAGCCGCTACCTGAGAGAGCTG 927
QY 751 GACAAAACAATAAGTTATACCTTACTTAACTTAACTGAGAGAG 795
DB 928 GACAAAGACCAAGCCGCTACTGAGCTACTTCTGCTGGCGGAG 972

RESULT 13

US-09-063-237-2
Sequence 2, Application US/09063237
Patent No. 6124267

GENERAL INFORMATION:

APPLICANT: McPeyer, Rodger P.
TITLE OF INVENTION: 0-Glycan Inhibitors of Selectin Mediated
TITLE OF INVENTION: Inflammation Derived from PSL-1
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30306-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/063,237
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/649,802
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRE110CIP7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-063-237-2

Query Match 12.3%; Score 112.2; DB 3; Length 2042;
Best Local Similarity 51.6%; Pred. No. 3e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;
QY 106 TGCCATCTCAACAGACGCTGCTGATACAAATCCCATGCACTTCTGATCCATCAC 165

DB 343 TGCCATCATATGCGCGACCGCAAGGTATACCCAGCAGCAACAGGTATCTGCGACAC 402
QY 166 CGAATACATCATTTGGG---TCGACAAATTTACCTCAGACACTGAGCAGCCCTTCGAG 222
DB 403 TGGATATCATATGTCACCACTTAAGTACAGCTCCACCTTCCCGAGGCGGAGGAG 462
QY 223 AATGATTTGATGAAATTTGGAATTCACCAATCAGACTCCCAAAAGAGTGGCATGAG 282
DB 463 CCGTGAATCTGTTCAACTTGAGACCCACCCCTACTGCGAGCAGTGAAGGCGCTGAGC 522
QY 283 CACTGTTTAACTGACTGACTTACCGCGGCTGATTCAGATTTCCAGTGCCTTAATGCG 342
DB 523 AGATACCTTCAATTCACCAATGCTCTACCGAGAGCTCCGACATCTTCAAGCCCTACGCG 582
QY 343 TTCTTGACGGTAAACCAAAATCCCTTCGT-----GTTTGAAGTCCCAAGCAAGAG 393
DB 583 TGCGTGAGCCGCTGTGCGGCAAGCTGCGCACCCACCCGCTCAACCTCTGCGCAAGAGC 642
QY 394 AATTTGCTGCTGCTGTTGTGATGATGAGTACGAGCCTGAGCATGCAAGATTTAC 453
DB 643 GAGCTGTGCGCTGGGCGGTGTCCAACTGGAAGCCGAGCTCAGCAGGGTGGCTACTAC 702
QY 454 AATGAGCTAACCAAAAGCATTTGAATTCATACCTACGCGGCAACCATTTGGAGATATGTC 513
DB 703 CAGAGCTGCAAGCTCATCTCAAGGTGAGCTGAGAGAGCTCCCAAGCCCTGCGCC 762
QY 514 AATGATAAAATTTGATTCCTACCATATCTGCTTGTAAATTTTATCTTCCCTTGAANA 573
DB 763 AAGGGAGACA--TGATGAGAGAGCTGTCCCGGTACAAAGTTCTACTGCGCTTCGAGAG 819
QY 574 TCAATCCAAAGATTTACATACAGGAAAGCTAT---ACAATGCTTTTGTGCTGCTCT 630
DB 820 TCCCTGACCCCGACTCATCTACACGAGAACGCTGAGAGAACCCCTGAGAGCGCTGAGGC 879
QY 631 GTACCTGTGTTCTGAGACCATCTAGGAGAAACTATGAGATTTATTCAGCAGATTC 690
DB 880 GTGCCCGGTGTGCTGGAGCCAGAGAGCAAGCACTAGAGAGGTCTGCGACCCGAGCGC 939
QY 691 TTCAATCATGTGAGAGATTTAACTCTCCAGAGAGCTGAGCAATATTCAGCAGAGTTC 750
DB 940 TTCAATCAGCTGAGAGACTTCCAGAGCCCAAGAGACTGCGCGGAGTCTGAGAGAGCTG 999
QY 751 GACAAAACAATAAGTTATACCTTACTTAACTTAACTGAGAGAG 795
DB 1000 GACAAAGACCAAGCCGCTACTGAGCTACTTCTGCTGGCGGAG 1044

RESULT 14

US-07-914-281-1
Sequence 1, Application US/07914281
Patent No. 5324663

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/914,281
 FILING DATE: 19920720
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2043 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: CDNA
 ANTI-SENSE: NO
 US-07-914-281-1

Query Match 12.3%; Score 112.2; DB 1; Length 2043;
 Best Local Similarity 51.6%; Pred. No. 3e-24;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY 106 TGGCATCTCACAAGCGAGCGTTCTACTGTACACAAATCCCATGCACTTGTGATCATCAC 165
 DB 343 TGGCATCTCACAAGCGAGCGTTCTACTGTACACAAATCCCATGCACTTGTGATCATCAC 402
 OY 166 CGAGCATAGTTGGGA---TCTGACAAATTTACCTCAGCAAGCTTGGCCAGCTTCCAG 222
 DB 403 TGGGATATATGTCCAAACCTTAAGTACGCTCCACCTTCCCGGCGGCGGAGGCGG 462
 OY 223 AATGATTTGGATGATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
 DB 463 CGGTGATCTGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 522
 OY 283 CACTGTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
 DB 523 AGATACCTCAATCTCAGCATGTCCTACGAGGAGTCCGACATCTTCAAGCCCTTACGG 582
 OY 343 TCTTGACGCTAAGCAACAATCCCTTCGT-----GTTTGAAGTGCACCAAGCAAGAG 393
 DB 583 TGGCTGAGACCGGTGTCGCGGCGGAGCGTCCACCCCTCAACCTCTCGGCGCAAGACC 642
 OY 394 AATGATGCTGCTGGTGTGAGTACGTAAGACCTGAGCATGCGTCAAGCATGATTTAC 453
 DB 643 GAGCTGTGCTGGTGTGAGTACGTAAGACCTGAGCATGCGTCAAGCATGATTTAC 702
 OY 454 AATGACCTAAGCAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 513
 DB 703 CAGAGCTCAGAGCTCATCTCAAGGTGAGCTGTACGAGGAGTCCGACCAAGCCCTGCGCC 762
 OY 514 AATGATTAATAATTTGATTCCTACCATATCTGCTGTAATTTTATCTTCTTGAATAT 573
 DB 763 AAGGGGACCA---TGATGAGAGCGCTGTCCGCTTACAGTTCCTGCTGCTCGAGAAC 819
 OY 574 TCAATCCACAGAGATTACATCAAGCAAGAAAGCTAT---ACAATGCTTTTCTGGTGGCTGT 630
 DB 820 TCTTTCACACCCGACTACATCAACGAGAGCTGTGAGGAGACCCCTGAGGCGCTGGGCC 879
 OY 631 GAGCTGTGTTTCTGGGACCATCTAGGAAAACTATGAGATTTATATTCACAGCATTTCA 690
 DB 880 GTGCCGCTGTGCTGGGCGCCGAGCAAGCACTACGAGAGGTCTCTGCGACCGCGAGCC 939
 OY 691 TTCAATTCATGCTGAGATTTATACCTTCCCATGAGCTTGGCAAGATATCTGAAGAGTCT 750
 DB 940 TTCTATCCAGCTGAGCATCTTCCAGAGACCCCAAGGAGCTGCGCTGCTTACCTGAGAGCTG 999
 OY 751 GACAAAACAAATTAATACCTTACTTACTTAACTTGAAGAG 795
 DB 1000 GACAAAGACACCGCCGCTACCTGAGCTACTTTCGCTGGCGGAG 1044

RESULT 15
 US-08-393-246-1
 ; Sequence 1, Application US/08393246
 ; Patent No. 5595900

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 TITLE OF INVENTION: GLYCOPOLYDIPS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994

APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2043 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: CDNA

ANTI-SENSE: NO
 US-08-393-246-1

Query Match 12.3%; Score 112.2; DB 1; Length 2043;
 Best Local Similarity 51.6%; Pred. No. 3e-24;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY 106 TGGCATCTCACAAGCGAGCGTTCTACTGTACACAAATCCCATGCACTTGTGATCATCAC 165
 DB 343 TGGCATCTCACAAGCGAGCGTTCTACTGTACACAAATCCCATGCACTTGTGATCATCAC 402
 OY 166 CGAGCATAGTTGGGA---TCTGACAAATTTACCTCAGCAAGCTTGGCCAGCTTCCAG 222
 DB 403 TGGGATATATGTCCAAACCTTAAGTACGCTCCACACTTCCCGGAGGCGGAGGCGG 462
 OY 223 AATGATTTGGATGATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
 DB 463 CGGTGATCTGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 522
 OY 283 CACTGTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
 DB 523 AGATACCTCAATCTCAGCATGTCCTACGAGGAGTCCGACATCTTCAAGCCCTTACGG 582
 OY 343 TCTTGACGCTAAGCAACAATCCCTTCGT-----GTTTGAAGTGCACCAAGCAAGAG 393

Tue Oct 8 10:18:29 2002

us-09-744-748-4_copy_1454_2368.rni

Page 12

```
Db 583 TGGCTGAGACCGCGTGGTCGGCCAGCCTGCCCCACCCACCGCTCAACCTCTGCGCCAGAGCC 642
OY 394 AAATTGGTGTGCTGGTGTGAGTAACCTGAGACCTGAGCANTGCCAGAGTCAAGTATTAC 453
Db 643 GAGCTGCTGGCTGGCGGCTGCTCAACTGGAACCGGACTCAGCCAGGGTGCCTACTAC 702
OY 454 AATGAGCTAAGCAAAAGCATTTGAATCCATACCTACGGCGAAGCATTGGAGATAATGTC 513
Db 703 CAGAGCCTGCAGGCTCATCTCAAGGTGAGCTGACGGACGCTCCACAGCCCCGTGCC 762
OY 514 AATGATAAAAATTTGATTCCTACCATATCTGTTGTAATTTATCTTCTTTGAAAT 573
Db 763 AAGGGGACCA---TGATGGAGACGCTGTCCGGTACAAAGTTTACCTGGGCTTCGAGAAC 819
OY 574 TCAATCCACAAGGATTACATCACGSAAGAAAGCTAT---ACAATGCTTTCTGCGCTGCTCT 630
Db 820 TCCTTGACACCCGACTACATCACCGAAGAGCTGTGAGAGAACGCCCTGGAGGCTGGGCC 879
OY 631 GTACCTGTGTCTGGGACCATCTAGGAAAAATATGAGAATTATATTCAGAGATTCA 690
Db 880 GTGCCCTGTGCTGGCGCCACAGACGACACTACGAGAGGTTCCCTGCCACCCGAGGCC 939
OY 691 TTCATTGATGTGAGAGATTATACTCTCCAGTGAAGTACCAAGTATCTGAAGAAATC 750
Db 940 TTCTATCCAGCTGAGCAGACTTCCAGAGCCCCCAAGGACCTGGCCGTTACCTGAGAGCTG 999
OY 751 GACAAAAACATTAAGTTTATTAAGTACTTACTTTAACTGGAGAG 795
Db 1000 GACAAAGGACCAAGCCCGCTACTGAGCTACTTTCGCTGGCGGGAG 1044
```

Search completed: October 6, 2002, 07:21:30
Job time : 28.6694 secs

QY	439	AAAAAGAAAACCTCTTTTCCACCAAAAGTATTTATTTATGAAACTACTATTTCTGTG	498
QY	439	AAAAAGAAAACCTCTTTTCCACCAAAAGTATTTATTTATGAAACTACTATTTCTGTG	498
Db	249	AAAAAGAAAACCTCTTTTCCACCAAAAGTATTTATTTATGAAACTACTATTTCTGTG	308
QY	499	TGGGTGTGGCCATTGGGCGAGACCTTTGACCTTACATCTCGCCAGCAAGATGTTCAACATC	558
QY	499	TGGGTGTGGCCATTGGGCGAGACCTTTGACCTTACATCTCGCCAGCAAGATGTTCAACATC	558
Db	309	TGGGTGTGGCCATTGGGCGAGACCTTTGACCTTACATCTCGCCAGCAAGATGTTCAACATC	368
QY	559	CAAGCATGCCATCTCCACAAAGCAACCTTCACTGTACAAACAAATCCCATGCAAGTTCTGATC	618
QY	559	CAAGCATGCCATCTCCACAAAGCAACCTTCACTGTACAAACAAATCCCATGCAAGTTCTGATC	618
Db	369	CAAGCATGCCATCTCCACAAAGCAACCTTCACTGTACAAACAAATCCCATGCAAGTTCTGATC	428
QY	619	CATCACCGAGCATCGAGTGGGATCTGACAAATTTACTTACGACAGCTAGCCACCTTC	678
QY	619	CATCACCGAGCATCGAGTGGGATCTGACAAATTTACTTACGACAGCTAGCCACCTTC	678
Db	429	CATCACCGAGCATCGAGTGGGATCTGACAAATTTACTTACGACAGCTAGCCACCTTC	488
QY	679	CAGAAATGATTTGGATTTGGAAATTTGGAAATCCACACACACTCCCAAAAGATGGCAT	738
QY	679	CAGAAATGATTTGGATTTGGAAATTTGGAAATTTGGAAATTTGGAAATTTGGAAATTTGG	738
Db	489	CAGAAATGATTTGGATTTGGAAATTTGGAAATTTGGAAATTTGGAAATTTGGAAATTTGG	548
QY	739	GAGCACTTTTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG	798
QY	739	GAGCACTTTTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG	798
Db	549	GAGCACTTTTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG	608
QY	799	GGCTTCTGACGGTAGCAAAATCCCTTCTGTGTTGAAGTGGCAAGCAAGAAATTTG	858
QY	799	GGCTTCTGACGGTAGCAAAATCCCTTCTGTGTTGAAGTGGCAAGCAAGAAATTTG	858
Db	609	GGCTTCTGACGGTAGCAAAATCCCTTCTGTGTTGAAGTGGCAAGCAAGAAATTTG	668
QY	859	GTGTGTGGT	918
QY	859	GTGTGTGGT	918
Db	669	GTGTGTGGT	728
QY	919	CTAAGCAAAAGCAATTTAAATTCCTACTCTACGGGCAAGCATTTGGAATTTGCAATGAT	978
QY	919	CTAAGCAAAAGCAATTTAAATTCCTACTCTACGGGCAAGCATTTGGAATTTGCAATGAT	978
Db	729	CTAAGCAAAAGCAATTTAAATTCCTACTCTACGGGCAAGCATTTGGAATTTGCAATGAT	788
QY	979	AAAAATTTATTCCTCCATTCCTGCTGTGTAATTTATTTCTTCCTTTGAAATTTCAATC	1038
QY	979	AAAAATTTATTCCTCCATTCCTGCTGTGTAATTTATTTCTTCCTTTGAAATTTCAATC	1038
Db	789	AAAAATTTATTCCTCCATTCCTGCTGTGTAATTTATTTCTTCCTTTGAAATTTCAATC	848
QY	1039	CACAAAGATTTACATTCACGAGAAAGATATACATATGCTTTTCTGGCTGGCTGTGACCTGT	1098
QY	1039	CACAAAGATTTACATTCACGAGAAAGATATACATATGCTTTTCTGGCTGGCTGTGACCTGT	1098
Db	849	CACAAAGATTTACATTCACGAGAAAGATATACATATGCTTTTCTGGCTGGCTGTGACCTGT	908
QY	1099	CTTCTGGGACCATTCGAGGAAACATATGAGAAATTTATTTCTCAGACAGATTCATTCATTCAT	1158
QY	1099	CTTCTGGGACCATTCGAGGAAACATATGAGAAATTTATTTCTCAGACAGATTCATTCATTCAT	1158
Db	909	CTTCTGGGACCATTCGAGGAAACATATGAGAAATTTATTTCTCAGACAGATTCATTCATTCAT	968
QY	1159	GTGGAAGATTTATTAACCTCCAGTGAGCTAGCAAAAGTATCTGAAAGAGCGCAAAATAC	1218
QY	1159	GTGGAAGATTTATTAACCTCCAGTGAGCTAGCAAAAGTATCTGAAAGAGCGCAAAATAC	1218
Db	969	GTGGAAGATTTATTAACCTCCAGTGAGCTAGCAAAAGTATCTGAAAGAGCGCAAAATAC	1028
QY	1219	AATTAAGTTTACTTACTTACTTAACTGTGAGAGAGATTTCACTGTAATTTCTCCACGA	1278
QY	1219	AATTAAGTTTACTTACTTACTTAACTGTGAGAGAGATTTCACTGTAATTTCTCCACGA	1278
Db	1029	AATTAAGTTTACTTACTTACTTAACTGTGAGAGAGATTTCACTGTAATTTCTCCACGA	1088
QY	1279	TTTTGGGATTCACATGCTGTTGGCTGTGAGATCTGCAAAAAGCATCAAGATATTAAG	1338
QY	1279	TTTTGGGATTCACATGCTGTTGGCTGTGAGATCTGCAAAAAGCATCAAGATATTAAG	1338
Db	1089	TTTTGGGATTCACATGCTGTTGGCTGTGAGATCTGCAAAAAGCATCAAGATATTAAG	1148
QY	1339	TCTGTGTGTAATTTAGAGAAATGGTTTTGGAATTTAAATTTTTCATCACTGTGCAACTTG	1398
QY	1339	TCTGTGTGTAATTTAGAGAAATGGTTTTGGAATTTAAATTTTTCATCACTGTGCAACTTG	1398
Db	1149	TCTGTGTGTAATTTAGAGAAATGGTTTTGGAATTTAAATTTTTCATCACTGTGCAACTTG	1208
QY	1399	ATTAATATTTTATGAGATATCTCCAGATATTTGAGAGATTAAGAAGATGCAACATACTA	1458
QY	1399	ATTAATATTTTATGAGATATCTCCAGATATTTGAGAGATTAAGAAGATGCAACATACTA	1458
Db	1209	ATTAATATTTTATGAGATATCTCCAGATATTTGAGAGATTAAGAAGATGCAACATACTA	1268
QY	1459	CTTTTGTGTACAAATTTATTTTATACCTCTCTAGGGTAGAAGATATTTTGTGGA	1518
QY	1459	CTTTTGTGTACAAATTTATTTTATACCTCTCTAGGGTAGAAGATATTTTGTGGA	1518
Db	1269	CTTTTGTGTACAAATTTATTTTATACCTCTCTAGGGTAGAAGATATTTTGTGGA	1328

OY	1519	GATTTTAAAAAGCTCACACATGAGCAAAATCAATCCATTCGGTTTTAAATTTCCGTGATATA	1578
Db	1329	GATTTTTAAAAAGCTCACACATGAGCAAAATCAATCCATTCGGTTTTAAATTTCCGTGATATA	1388
OY	1579	CCATAATATGCGCATGAGAGATTAATTTATTCCTCATTTATCATTTTGTATAACATTCCTTTT	1638
Db	1389	CCATAATATGCGCATGAGAGATTAATTTATTCCTCATTTATCATTTTGTATAACATTCCTTTT	1448
OY	1639	TCACATTTTTGCTAGTTTCCATTAATGTAAGCTTGTGGTTGATATTATTTTCCACACGTGA	1698
Db	1449	TCACATTTTTGCTAGTTTCCATTAATGTAAGCTTGTGGTTGATATTATTTTCCACACGTGA	1508
OY	1699	TCAGCTGTTTAACTATTGTTGGAAATGAAATGATGACATCTTAAAGATGAAAAATTTTCA	1758
Db	1509	TCAGCTGTTTAACTATTGTTGGAAATGAAATGATGACATCTTAAAGATGAAAAATTTTCA	1568
OY	1759	CTAAGATATTACATGCTCTAGTTCCAACTTTGCAATACATTAACAGAGAGAACAATGTTGC	1818
Db	1559	CTAAGATATTACATGCTCTAGTTCCAACTTTGCAATACATTAACAGAGAGAACAATGTTGC	1628
OY	1819	GATTGGAATTTCAACCTCTTTTGACTCCTAAGATGAATGAAGTGTACTGTCTCATTTTG	1878
Db	1629	GATTGGAATTTCAACCTCTTTTGACTCCTAAGATGAATGAAGTGTACTGTCTCTGA-ITG	1687
OY	1879	ATCATTTTTTTTTTACCGTTTATTCACATTTTGTGAAGTGAATTTATCATGAGTGATA	1938
Db	1688	ATCATTTTTTTTTTACCGTTTATTCACATTTTGTGAAGTGAATTTATCATGAGTGATA	1747
OY	1939	AGAAAGATGACAGACACACATGCTTATTCAGATTCAGACACCTTATAGCTCTCATTTATTC	1998
Db	1748	AGAAAGATGACACACACATGCTTATTCAGATTCAGACACCTTATAGCTCTCATTTATTC	1807
OY	1999	ATTAAAGCATTTGACACTACTATTATTCATCGATCGTCTAAATTAATTAATATCAAGATTT	2058
Db	1808	ATTAAAGCATTTGACACTACTATTATTCATCGATCGTCTAAATTAATTAATATCAAGATTT	1867
OY	2059	TAAATATTCACATTTGTTGTTGATTTTTCAGACACCGGGAGATATATCCAAATATACCTTGA	2118
Db	1868	TAAATATTCACATTTGTTGTTGATTTTTCAGACACCGGGAGATATATATCCAAATATACCTTGA	1928
OY	2119	AAATCTTAAGACAGTTCTTTGCGTACAGATGACATCATTTGCAATATATTAACAAATAT	2178
Db	1927	AAATCTTAAGACAGTTCTTTGCGTACAGATGACATCATTTGCAATATATTAACAAATAT	1986
OY	2179	TTTCCGCAAAATTAACAAACAAAATGATATCCATTAATATTTATTAATAGTGTCAATTTAT	2238
Db	1987	TTTCCGCAAAATTAACAAACAAAATGATATCCATTAATATTTATTAATAGTGTCAATTTAT	2046
OY	2239	GAAACAATGTTTAATTAATGATGTCATATTTAAGATTTTTCGAAAGCCCTAATATTTAAAT	2298
Db	2047	GAAACAATGTTTAATTAATGATGTCATATTTAAGATTTTTCGAAAGCCCTAATATTTAAAT	2106
OY	2299	GGCCTTATTTTACCATATGATATTAAGATTTGGCTCATATGATGAGCCCTATCATTTGA	2358
Db	2107	GGCCTTATTTTACCATATGATATTAAGATTTGGCTCATATGATGAG-CCATATCTTTGA	2165
OY	2359	TTTACAGTCTCTCATTTTAAGAGAGCCCAATTAATTAATTAATCATCAAGGTATTAAATTAAG	2418
Db	2166	TTTACAGTCTCTCATTTTAAGAGAGCCCAATTAATTAATTAATTAATCATCAAGGTATTAAATTAAG	2225
OY	2419	ACGTTAAATATAAATAAAGTGGAGATATATAGAAACAACACAGCTGTATGACACAGATTAAGA	2478
Db	2226	ACGTTAAATATAAATAAAGTGGAGATATATAG-AAAACAACAGCTGTATGACACAGATTAAGA	2288
OY	2479	TCCTAAATGACATTTGTTGATGATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2538
Db	2285	TCCTAAATGACATTTGTTGATGATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2344
OY	2539	GTTTCAAGGAAGTACACAGTCTCTACTATTAGAAAGTACATTAATTTGAGATGAATTTATATCC	2598
Db	2345	GTTTCAAGGAAGTACACAGTCTCTACTATTAGAAAGTACATTAATTTGAGATGAATTTATATCC	2404
OY	2599	ATTGTGAAATTTATTCATATATGACATATGTCACATTTGATTTGACAGGGCTTTATGAAAC	2658

|||||
Db 2405 ATTGGTATATTCATACATGACATATGATGTCAGGGCTTATAGAAC 2464
OY 2659 AAATAACCTACCATTTAC 2676
|||||
Db 2465 AAATAACCTACCATTTAC 2482
RESULT 3
AL512406
LOCUS
DEFINITION Human DNA sequence from clone RP11-504J9 on chromosome 6, complete
sequence.
ACCESSION AL512406
VERSION AL512406.14 GI:13897154
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 80247)
AUTHORS Bates, K.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Apr 30, 2001 this sequence version replaced gi:13396709.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
validation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw, S-
WISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-504J9 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-504J9 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP3-381A2 is at 80148 in this sequence.
The true right end of clone RP11-77B15 is at 100 in this sequence.
location/Qualifiers
1. 80247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-504J9"
/clone_lib="RPCI-11.2"
184. 1422
repeat_region
/note="L2 repeat: matches 875. .2419 of consensus"
1744. 1834
repeat_region
/note="MIR repeat: matches 166. .258 of consensus"
2950. 2977
repeat_region
/note="14 copies 2 mer ca 100% conserved"
3411. 3669
/note="LTR16C repeat: matches 93. .364 of consensus"

repeat_region
4372. 4704
/note="LTR1A1 repeat: matches 5. .365 of consensus"
misc_feature
complement(7821..8406)
/note="match: GSS: Em:AQ550748"
8126. 8184
repeat_region
/note="MER58B repeat: matches 272. .332 of consensus"
8185. 8696
repeat_region
/note="LTR2D repeat: matches 1. .513 of consensus"
8779. 8817
repeat_region
/note="LTR2D repeat: matches 504. .542 of consensus"
8819. 8974
repeat_region
/note="MER58A repeat: matches 6. .162 of consensus"
complement(9443..9891)
/note="match: GSS: Em:B90633"
complement(9452..9883)
/note="match: GSS: Em:AQ119695"
9782. 9894
repeat_region
/note="MIR repeat: matches 92. .205 of consensus"
10089. 10212
repeat_region
/note="MIR repeat: matches 135. .262 of consensus"
complement(10093..10671)
/note="match: GSS: Em:AQ237061"
10223. 10282
repeat_region
/note="MIR repeat: matches 203. .262 of consensus"
complement(10315..10631)
/note="match: GSS: Em:AQ200259"
10805. 11186
misc_feature
/note="match: GSS: Em:AQ072353"
11657. 11844
repeat_region
/note="L2 repeat: matches 2546. .2750 of consensus"
complement(11659..12084)
/note="match: GSS: Em:AQ0518361"
complement(11674..11987)
/note="match: GSS: Em:AQ094678"
12083. 12588
misc_feature
/note="match: GSS: Em:AQ691235"
12596. 12810
repeat_region
/note="MIR repeat: matches 23. .248 of consensus"
12920. 12986
repeat_region
/note="L1M45 repeat: matches 6228. .6295 of consensus"
12987. 13169
repeat_region
/note="L1M45 repeat: matches 5852. .6043 of consensus"
13168. 13280
repeat_region
/note="L1M48 repeat: matches 5710. .5832 of consensus"
13851. 17218
repeat_region
/note="L1P1 repeat: matches 2787. .6155 of consensus"
18903. 18953
repeat_region
/note="LTR2CB repeat: matches 439. .501 of consensus"
18984. 19433
repeat_region
/note="LTR2CB repeat: matches 1. .460 of consensus"
complement(19108..19690)
/note="match: GSS: Em:AQ563935"
20683. 20801
repeat_region
/note="AluYo/FRAM repeat: matches 169. .288 of consensus"
20802. 20862
repeat_region
/note="AluYo/FRAM repeat: matches 3. .66 of consensus"
20912. 21033
repeat_region
/note="MIR repeat: matches 114. .237 of consensus"
21047. 21128
repeat_region
/note="41 copies 2 mer ta 63% conserved"
21676. 21859
repeat_region
/note="L2 repeat: matches 2130. .2308 of consensus"
21890. 23612
repeat_region
/note="L1P repeat: matches 841. .2570 of consensus"
23601. 26609
repeat_region
/note="L1P45 repeat: matches 3107. .6145 of consensus"
26620. 26789
repeat_region
/note="MER5A repeat: matches 1. .174 of consensus"
28521. 28627
repeat_region
/note="MER39 repeat: matches 5. .111 of consensus"
28656. 28837
repeat_region
/note="MER39b repeat: matches 54. .235 of consensus"
28838. 29142

```
repeatL_region /note="AluSc repeat: matches 4. .308 of consensus"  
                29143. .29494  
                /note="MER39b repeat: matches 235. .549 of consensus"  
                30529. .31344  
                /note="LiP43 repeat: matches 5331. .6146 of consensus"  
                32234. .32340  
                /note="L12 repeat: matches 2627. .2750 of consensus"  
                32399. .32578  
                /note="MIR repeat: matches 12. .203 of consensus"  
                32816. .32979  
                /note="MIR repeat: matches 91. .259 of consensus"  
                complement(33006. .33450)  
                /note="match: GSS: Em:AQ261227"  
                complement(33058. .33444)  
                /note="match: GSS: Em:AQ404053"  
                33312. .33596  
                /note="LM4 repeat: matches 3473. .3773 of consensus"  
                34856. .34915  
                /note="30 copies 2 mer aa 68% conserved"  
                complement(34987. .35547)  
                /note="match: GSS: Em:AQ321598  
                match: STS: Em:G55571"  
                complement(36488. .36936)  
                /note="match: GSS: Em:B41073"  
                complement(36654. .37090)  
                /note="match: GSS: Em:AQ772584"  
                37102. .37591  
                /note="match: GSS: Em:AQ489567"  
                complement(37581. .38221)  
                /note="match: GSS: Em:AQ418233"  
                37720. .37772  
                /note="L12 repeat: matches 2657. .2710 of consensus"  
                complement(37778. .38215)  
                /note="match: GSS: Em:AQ819773"  
                38218. .38732  
                /note="match: GSS: Em:AQ599868"  
                38861. .39164  
                /note="AluXc repeat: matches 1. .305 of consensus"  
                39612. .39653  
                /note="21 copies 2 mer ac 90% conserved"  
                40854. .40909  
                /note="L12 repeat: matches 2655. .2710 of consensus"  
                42169. .42270  
                /note="MER5A repeat: matches 1. .111 of consensus"  
                42411. .42548  
                /note="MER5B repeat: matches 21. .181 of consensus"  
                42780. .42938  
                /note="MER5B repeat: matches 17. .177 of consensus"  
                43229. .43326  
                /note="MIR repeat: matches 40. .148 of consensus"  
                45383. .45428  
                /note="MER5B repeat: matches 18. .62 of consensus"  
                complement(47284. .47971)  
                /note="match: STS: Em:C51995  
                match: GSS: Em:AQ269059"  
                complement(47537. .48011)  
                /note="match: STS: Em:G36899"  
                complement(47723. .48014)  
                /note="match: STS: Em:G30049"  
                47972. .48451  
                /note="match: GSS: Em:AQ336527"  
                complement(48518. .48985)  
                /note="match: GSS: Em:AQ801834"  
                48597. .48803  
                repeatL_region
```

	Query Match:	89.0%	Score 2380.4	DB 9	Length 80247
	Best Local Similarity	99.7%	Pred 10.0		
	Matches 2384	Conservative	0	Mismatches 6	Indels 0
				Gaps	0
QY	287	GAATAATATGACATCAATCAATCAAGAAATCTCGCCATTTTAAATGTCGCATTA	346		
Db	35551	GAATAATATGACATCAATCAATCAAGAAATCTCGCCATTTTAAATGTCGCATTA	35610		

QY	347	TCCTGGCTGTTTCAATGAGCATGTCTTTCATTTACATCAACCCACCAACAGCTGGATCT	406
Db	35611	TCCTGGCTGTTTCAATGAGCATGTCTTTCATTTACATCAACCCACCAACAGCTGGATCT	35678
QY	407	TCAGTCCAAATGGAATCAAGCCAGCTCTGTGCTGAAAAATGAAAAACTTCTTTCACACAAA	466
Db	35671	TCAGTCCAAATGGAATCAAGCCAGCTCTGTGCTGAAAAATGAAAAACTTCTTTCACACAAA	35737
QY	467	CTGATTATTTTAAAGAAACACTAATTTCTGTGTGGGTGTGGCCATTGTGGGAGACCTTGG	526
Db	35731	CTGATTATTTTAAAGAAACACTAATTTCTGTGTGGGTGTGGCCATTGTGGGAGACCTTGG	35799
QY	527	ACCTTACATTCCTGGCAAGCAATGTTCAACATCCCAAGATGCCATCTCACACAGGACCGTT	586
Db	35791	ACCTTACATTCCTGGCAAGCAATGTTCAACATCCCAAGATGCCATCTCACACAGGACCGTT	35858
QY	587	CACCTACACAACAAATCCCATGTCATGTGTATCCATCACAGAGCATCACTTGGATCTGGA	646
Db	35851	CACCTACACAACAAATCCCATGTCATGTGTATCCATCACAGAGCATCACTTGGATCTGGA	35911
QY	647	CAAAATTTACCTCAGACAGCTAAGGACACCTTCCAGAAATGAGATTSAGATTTTGGAT	706
Db	35911	CAAAATTTACCTCAGACAGCTAAGGACACCTTCCAGAAATGAGATTTTGGATTTTGGAT	35971
QY	707	CACCAACATCACACATCCCCAAAAAGGTGGCATTTGAGCACTTGTTTAACTCTGACTCTGACT	766
Db	35971	CACCAACATCACACATCCCCAAAAAGGTGGCATTTGAGCACTTGTTTAACTCTGACTCTGACT	36031
QY	767	ACCGCCGTGATTTCAGATATCCAGATGCGCTTATGGCTTCTTGACGGTACACAAATCCCT	826
Db	36031	ACCGCCGTGATTTCAGATATCCAGATGCGCTTATGGCTTCTTGACGGTACACAAATCCCT	36091
QY	827	TCGTGTTGGAAGTGCACAAAGCAAGAAATTTGGTGTGCTGGGTGTGAGTAATCTGGAAC	886
Db	36091	TCGTGTTGGAAGTGCACAAAGCAAGAAATTTGGTGTGCTGGGTGTGAGTAATCTGGAAC	36151
QY	887	CTGAGCATGCCAGAGTCAAGATTTACAAATAGCTAAGCAACAAAGCATTTGAAATCCATACT	946
Db	36151	CTGAGCATGCCAGAGTCAAGATTTACAAATAGCTAAGCAACAAAGCATTTGAAATCCATACT	36211
QY	947	ACGGGCAAGCAATTTGGAGAAATATCTCAATGATAAAAATTTGGATTCCTTACATATCTGCT	1006
Db	36211	ACGGGCAAGCAATTTGGAGAAATATCTCAATGATAAAAATTTGGATTCCTTACATATCTGCT	36271
QY	1007	GTAATTTTATCTTCCCTTTSAAATATCATATCCACAAGGATTCATCATCAGGAAAAAGCTAT	1066
Db	36271	GTAATTTTATCTTCCCTTTSAAATATCATATCCACAAGGATTCATCATCAGGAAAAAGCTAT	36331
QY	1067	ACAAATGCTTTCTTGGGCGCTGTGTACTCTTGTGTGGGACCATCTAAGGAAAACTATG	1126
Db	36331	ACAAATGCTTTCTTGGGCGCTGTGTACTCTTGTGTGGGACCATCTAAGGAAAACTATG	36391
QY	1127	AGAAATTAATTCGAGAGATTCATCTCATCTCATGTGAGAAATTTAACTCTCCCACTAGC	1186
Db	36391	AGAAATTAATTCGAGAGATTCATCTCATCTCATGTGAGAAATTTAACTCTCCCACTAGC	36451
QY	1187	TAGCAAAATATCTGAGAGAGTGCACAAAACAATAAGTATACCTTAGTACTTTAACT	1246
Db	36451	TAGCAAAATATCTGAGAGAGTGCACAAAACAATAAGTATACCTTAGTACTTTAACT	36511
QY	1247	GGAGGAGGATTTCACTGTAAATCTTCCACGATTTTGGGAATCACATGCAATTTGGCTT	1306
Db	36511	GGAGGAGGATTTCACTGTAAATCTTCCACGATTTTGGGAATCACATGCAATTTGGCTT	36571
QY	1307	GGGATCATGTAAAAAGGCATTCAGAAATATATAGTCTGTGGTAAATTTAGAAATGGTTTT	1366
Db	36571	GGGATCATGTAAAAAGGCATTCAGAAATATATAGTCTGTGGTAAATTTAGAAATGGTTTT	36631
QY	1367	GGAAATTAATTTTTCATCTGACATTTGCACACTGTGATTAATTTTGTAGAGAAATCATCCAA	1426
Db	36631	GGAAATTAATTTTTCATCTGACATTTGCACACTGTGATTAATTTTGTAGAGAAATCATCCAA	36691
QY	1427	GTATTGAGGATTAAGAGATGCAACATCACTTCTTGTGTGACAAATTAATTTATATCAC	1486

C 1486

```

|||||
Db 36691 GATTAGAGTAAAGAGAGAGACATCTACTTTTGTGTCACAAATTTATTTTATCAG 36750
|||||
QY 1487 CCTCTAGAGGTAAAGGTATATTTTGTGAGATTTTAAAGCTCAGCATGAGCAATC 1546
|||||
Db 36751 CCTCTAGAGGTAAAGGTATATTTTGTGAGATTTTAAAGCTCAGCATGAGCAATC 36810
|||||
QY 1547 ATTCCATTCGGTTTAAATTAATTCCTGTATATACCTAATTAATGTCAGCTGAGATATTT 1606
|||||
Db 36811 ATTCCATTCGGTTTAAATTAATTCCTGTATATACCTAATTAATGTCAGCTGAGATATTT 36870
|||||
QY 1607 ATTCTCATTAATACATTTGTAACATGCTTTTACATTTTGTAGTGTCCATTAATGTA 1666
|||||
Db 36871 ATTCTCATTAATACATTTGTAACATGCTTTTACATTTTGTAGTGTCCATTAATGTA 36930
|||||
QY 1667 AGCTTGTGGTTGATTAATTTTCCACACATGATCAGCTGTTTATGTAATTTGGGAAATGA 1726
|||||
Db 36931 AGCTTGTGGTTGATTAATTTTCCACACATGATCAGCTGTTTATGTAATTTGGGAAATGA 36990
|||||
QY 1727 AGATGCAATCTTAAGATGAATAATTTTCACTAAGATTAATGATGCTAGTTCCAACT 1786
|||||
Db 36991 AGATGCAATCTTAAGATGAATAATTTTCACTAAGATTAATGATGCTAGTTCCAACT 37050
|||||
QY 1787 TTGCATACATTAATACAGAGAAAGACATGTCGATGAATTTCACTGCTTACCTCCTTA 1846
|||||
Db 37051 TTGCATACATTAATACAGAGAAAGACATGTCGATGAATTTCACTGCTTACCTCCTTA 37110
|||||
QY 1847 AGATGAATGAAGTGTAACTGCTCTATATTTGATGCTATTTTATTTTACCTGTTATACAT 1906
|||||
Db 37111 AGATGAATGAAGTGTAACTGCTCTATATTTGATGCTATTTTATTTTACCTGTTATACAT 37170
|||||
QY 1907 TTGTGAAGTGAATTAATTAATTCAGAGTGAATTAAGAAAGATATGACAGACAGCTGTTCTAT 1966
|||||
Db 37171 TTGTGAAGTGAATTAATTAATTCAGAGTGAATTAAGAAAGATATGACAGACAGCTGTTCTAT 37230
|||||
QY 1967 TCAGGAAGCATATAGACTCTCATTTATTTTCAATTAAGCTGATTTGACAGCTTATCTT 2026
|||||
Db 37231 TCAGGAAGCATATAGACTCTCATTTATTTTCAATTAAGCTGATTTGACAGCTTATCTT 37290
|||||
QY 2027 CAGGTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2086
|||||
Db 37291 CAGGTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37350
|||||
QY 2087 CACCTGGGAAGTAATCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2146
|||||
Db 37351 CACCTGGGAAGTAATCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37410
|||||
QY 2147 ATGACACTCATTTGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2206
|||||
Db 37411 ATGACACTCATTTGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37470
|||||
QY 2207 CTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2266
|||||
Db 37471 CTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37530
|||||
QY 2267 AGATTTTTTTCTGAAGCCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2326
|||||
Db 37531 AGATTTTTTTCTGAAGCCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37590
|||||
QY 2327 TTGGGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2386
|||||
Db 37591 TTGGGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37650
|||||
QY 2387 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2446
|||||
Db 37651 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 37710
|||||
QY 2447 TAGAAAAACACAGAGTTTACACAGAGTAAGATCTCAATGACCAATTTGTTGATGAATAA 2506
|||||
Db 37711 TAGAAAAACACAGAGTTTACACAGAGTAAGATCTCAATGACCAATTTGTTGATGAATAA 37770
|||||
QY 2507 ATTAATGCAATTTGATTTCCAGAAAAATGATTTGTTTCAAGGAAGTGACAGTTCTACTTTA 2566
|||||

```

```

Db 37771 ATTAATGCAATTTGATTTCCAGAAAAATGATTTGTTTCAAGGAAGTGACAGTTCTACTTTA 37830
QY 2567 GAAGTACTAATTTGGAGATGACCTTTATATCCATTTTGTGTAATTAATTAATTAATTAATTAATTAAT 2626
Db 37831 GAAGTACTAATTTGGAGATGACCTTTATATCCATTTTGTGTAATTAATTAATTAATTAATTAATTAAT 37890
QY 2627 TATGACCATGATGTTTCAGGCTTTATAGAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2676
Db 37891 TATGACCATGATGTTTCAGGCTTTATAGAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 37940

RESULT 4
AB015426 2139 bp mRNA linear ROD 23-JUN-1999
LOCUS Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.
ACCESSION AB015426
VERSION AB015426.1 GI:3702718
KEYWORDS Fut9, alpha1,3-fucosyltransferase IX.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Kudo,T., Ikehara,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K. and Narimatsu,H.
TITLE Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, fMuc-RIX, that synthesizes the Lewis x (CD15) epitope in brain and kidney
J. Biol. Chem. 273 (41), 26729-26738 (1998)
98434588
JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 2139)
AUTHORS Kudo,T. and Narimatsu,H.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science, Soka University, Division of Cell Biology; 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan (E-mail:tkudo@et.soka.ac.jp, tel: +81-426-91-2495, Fax: +81-426-91-9315)
Location/Qualifiers
1. 2139
/organism="Mus musculus"
/db_xref="taxon:10090"
115..1194
/gene="Fut9"
115..1194
/gene="Fut9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="BAA33522.1"
/db_xref="GI:3702719"
/translation="MTSTSKGILRPFLIVCIILGCFMACLLIYIRPTNSWVSPMESA
SSVLKMKNFESTKTDVFNFTLLVWVPGQTFDLSQCAMINISGCHLITPRLSYLNK
SHAVLILHRDIDSLMDLNLPOARPPOKRIWNLSPPRPOKSGEHLFNTLYRKR
BDDIQPYPLFVYSTRPVEYEPVSKELVGYVSNPNPAAKRYINELSKSIEHTY
GAFGEYVNDKMLIPTISTCKFYLSFENSIRHVDYITERKLYNAFLAGSVVPLGPKSN
YENYIPADSPFIHVEDPNSPELAKYLEKVDKNNKLYLSYFNMKDEFTVNLPRWESHA
CLACDHVKRHOEKKSYGNLEKFWN"
2139
polya_site
/note="31 a nucleotides"
BASE COUNT 684 a 395 c 368 g 692 t
ORIGIN
Query Match 45.1%; Score 1206.8; DB 10; Length 2139;
Best Local Similarity 78.4%; Pred. No. 1.1e-233;
Matches 1674; Conservative 0; Mismatches 387; Indels 75; Gaps 16;
QY 187 CCTTAGGACCGATTGTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 246
Db 9 CGCTGGGATGCTTTAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 68
QY 247 ACAACGTCTACGTCGCTCCCATGATATGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 306
Db 69 GCACTCATCTC-TGCTTCCATGCTATGTTCTTACACGAAAAAATTAATGACATCAACA 126

```

OY	307	TCCAAAGAAATTCCTGCGCCATTTTAAATGTCGCAATATCCTGGGCGTTCCATGCGA	366
Db	127	TCCAAAGGCATTTCCTGCGCCATTTCTAATGCTGTGATCATCTGGGCGTTCCATGCGA	186
OY	367	TGCTCTGTCATTTCATCAAAACCTACCAACAGCTGGATCTTCAGCTCCANTGAAATCAGC	426
Db	187	TGCTCGTCATTTCATCAAGCCACCACAAGCTGGGTCTCAGTCCAAATGCAAGTGTGCA	246
OY	427	AGCTGTGGCGAAATGAAAACCTGCTTCCACCAAAACGATATTTAAATGAAAT	486
Db	247	AGTTCTGTGCGAATAATGAAAATTTCTTCCACAAAACGTATTTATTTAACGAAACT	306
OY	487	ACTATCTGGGTGGGTGTGCGCCATTTGGGCAACCTTTAGCTTCATCTCCACAGA	546
Db	307	ACCATCTGGGTTTGGGTATGGCCATTGGGCAAGCCTTTGACTTACATCTCTGCCAAGA	366
OY	547	AWGTTCAACATCCAAAGATGCCATCTCCAAAGGACCGTTCACTGTACACAACAAATCCAT	606
Db	367	AWGTTCAATATCCAAAGGTGCCATCTCCAAAGACCGCGTCAATTTGACACAACAAATCCAT	426
OY	607	GCAGTTCTGATCCATCCACCGAGCATAGTTGGGACTGTGACAAATTTACCTACGCAAGCT	666
Db	427	GCGGCTTCGATTCACCATAGAGCATCCAGTGGGATCTGTACTAACTTACCTACGCAAGCC	486
OY	667	AGGCGACCCCTCCGAAATGATTTGGATTAATTTGGAAATCACCAACTCACACTCCCA	726
Db	487	AGGCGACCCCTTCGAAATGATTTGGATTAATTTAGAGTACCCCTCACACCCCCCA	546
OY	727	AAGAGTGCAATTGAGCACTGTTTAACTCTGACTGTGACTACCGCGGTACATCAGATATC	786
Db	547	AAGATGGCAATTGAACACTGTTTCAACTCTGACTTAATCTTATCGCGGTACAGATATC	606
OY	787	CAGAAGCCTTTAGCCTCTTGAGCGGTAGACAAATCCCTGTGTTTAAAGTGGCAAC	846
Db	607	CAAGTGCTTATGGCTCTTGAGCGGTAGACAAATCCCTGTGTTTAAAGTGGCAAC	666
OY	847	AAAGGAAATTTGGTGTGCTGGTTGTAGTAACTGGAAACCTCGAGATCCAGAGTCCAG	906
Db	667	AAGGAAAGTTGGTGTGCTGGTTGTAGTAACTGGAAACCTCGAGATCCAGAGTCCAG	726
OY	907	TATTACATATGCTTAACCAAAAGCATTTGAAATCCATACCTACGGGCAACCATTTGGAGA	966
Db	727	TATTACAGAGCTCCAGCAAGATATGAATCCACACTGATGGCCACAGCATTTGGAGA	786
OY	967	TATGTCAATGATTAATAATTTGATTTCTTACCATATCTGCTGTGTAATTTATCTTCTCTT	1026
Db	787	TACGTGAACGATTAATAATCTGATTTCCACCATATCTACTGTGTAATTTATCTTCTCATTT	846
OY	1027	GAAATATCAATCCAAAGATATTCATCCAGGAAAGCATATACAAATGCTTTCTGGCTGTC	1086
Db	847	GAAATATCAATTCACAAAGATATTCATCCAGGAAAGCTCTACAAATGCTATTTTGGCTGAT	906
OY	1087	TCTGTACCTGTGTTTTCGGGACCATCTAGGGAAGACATATAGAAATATATTTCCAGACAT	1146
Db	907	TCAGTACCTGTGTTCTCGGGGTCCATCTAGGGAAGACATATAGAAATATATTTCCAGCTGAT	966
OY	1147	TCATTCATTCATGTGGAAAGATTTAATCTCCAGAGAGCTACACAAGATATCGAAGAA	1206
Db	967	TCATTCATTCATGTGGAAAGATTTAATCTCCAGAGAGCTATACAAAATATCGAAGAA	1026
OY	1207	GTCGCAAAAACATTAGTATATACCTATGACTTTTAACTGCGGAGGAAAGATTTCCACTGTA	1266
Db	1027	GTTGCAAAAACATTAGTATGTTGACTTATGTTTAACTGCGGAGAAAGATTTTACTGTA	1086
OY	1267	AATCTCCAGATTTTGGGAAATCAACATGATCTTTGGCTTTGGAGATCTGAAAAGCAT	1326
Db	1087	AACCTACAGCGTTTGGGATTCACATGATCTGCTGCATCGACATCTTAAAAGGCAT	1144
OY	1327	CAAGATATATAGTCTGTGGTAAATTTAGAGAAATGTTTGGATTAATAATTTTTCATCA	1386
Db	1147	CAAGATATATAGTCTGTGGTAAATTTAGAGAAATGTTTGGATTAATAATTTTTCATCA	1206

[illegible]

SOURCE Ratius norvegicus cdna to mRNA.
 ORGANISM Ratius norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Ratius.
 REFERENCE 1 (sites)
 AUTHORS Shimoda,Y., Tajima,Y., Osanai,T., Katsume,A., Kohara,M., Kudo,T.,
 Narimatsu,H., Takashima,N., Ishii,Y., Nakamura,S., Osumi,N. and
 Sanai,Y.
 TITLE Pax6 Controls the Expression of Lewis X Epitope in the Embryonic
 Forebrain by Regulating alpha 1,3-Fucosyltransferase IX Expression
 JOURNAL J Biol. Chem. 277 (3), 2033-2039 (2002)
 PUBMED 11675393
 REFERENCE 2 (bases 1 to 2156)
 AUTHORS Sanai,Y.
 TITLE Direct Submission
 JOURNAL Submitted (11-OCT-2000) Yutaka Sanai, Tokyo Metropolitan Institute
 of Medical Science, Department of Biochemical Cell Research;
 Honkomagome 3-18-22, Bunkyo-Ku, Tokyo 113-8613, Japan
 (E-mail:sanaie@insoken.or.jp, Tel:81-3-3823-2101(ex.5233),
 Fax:81-3-3828-6663)
 FEATURES
 source location/Qualifiers
 1. 2156
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 gene 346..1425
 /gene="FUT9"
 CDS 346..1425
 /gene="FUT9"
 /codon_start=1
 /product="alpha1,3-fucosyltransferase IX"
 /protein_id="BAB40953.1"
 /db_xref="GI:13591589"
 /translation="MSTSKGLRPLIVLCIIIGCFVACLLIYIKPTNSWVSPMESA
 SSVLMKNFESKTYENETITLWVWPGTEFDLSCAMNIOGCHLTPRSLYRK
 SHAVLIHRDIDISMDLNLPOQARPPQKMTIMNBSPTPQSGIEHLENILYTR
 BSDIDVPGFIVLSTSPFVEVPSKELKLVWVSNPBARKKYINELSKSEIHTY
 GQAFGEVNDKMLIPTISCKEYLSFENSIHDKYIEKLYNFAIGSVLVGPSHEN
 YENYIPADSFIVHEIDNSPELAKYLYKEVDKNNKLYLSYFNMRKDETVNLPRWESHA
 CLACDHVRKHQEKYSVGNLEKMFVN"
 BASE COUNT 605 a 470 c 437 g 644 t
 ORIGIN
 Query Match 44.3%; Score 1184.8; DB 10; Length 2156;
 Best Local Similarity 78.2%; Pred. No.3.2e-229;
 Matches 1591; Conservative 0; Mismatches 392; Indels 51; Gaps 12;

Db 464 CAATGAGTCGCAAGCTCTGTCTGTAATAATGAAAAATTTCTCTCCAGGAAACGATT 523
 QY 473 ATTTAATGAAGTACTATTCGTGTGGTGTGGCCATTTGGGACAGACCTTGAACCTTA 532
 Db 524 ATTTAATGAAGTACTATTCGTGTGGTGTGGCCATTTGGGACAGACCTTGAACCTTA 583
 QY 533 CATCTGCGCAAGCAATGTTCAATCCAGAGATGCCATCTCACAGGACCGTTCACTGT 552
 Db 584 CATCTGCGCAAGCAATGTTCAATCCAGAGATGCCATCTCACAGGACCGTTCACTGT 643
 QY 593 ACAACAAATCCATGAGTCTGATCATCAACGAGACATCACTGGATCTACAAAT 652
 Db 644 ACAACAAATCCATGAGTCTGATCATCAACGAGACATCACTGGATCTACAAAT 703
 QY 653 TACCTGAGCAAGCTAGGCGCACCTTCAGAAATGATTTGGATGAATTCACCA 712
 Db 704 TACCTGAGCAAGGCGCGCCCTTCAGAAATGATTTGGATGAATTCACCA 763
 QY 713 CTCACACTCCCAAAAGATGCGATTGACGACCTTTTAACCTGACTCTGATCCGCC 772
 Db 764 CTCACACCCCAAAAGATGCGATTGACGACCTTTCAACCTGACTCTGATTCGCC 823
 QY 773 GTGATTCAGATTCAGAGTCCCTTAGGCTTCTGACGGTAAGCAAAATCCCTGTGT 832
 Db 824 GTGATTCAGATTCAGAGTCCCTTAGGCTTCTGACGGTAAGCAAAATCCCTGTGT 883
 QY 833 TTGAAGTCCCAAGCAAGAAATGGTGTGTGTGGTGTGATTAAGTCCAGACCTGAGC 892
 Db 884 TTGAAGTCCCAAGCAAGAAATGGTGTGTGTGGTGTGATTAAGTCCAGACCTGAGC 943
 QY 893 ATGCCAGATCAAGTATTACAAATGAGCTAAGCAAAAGCATTAATCCATACCTAGCGC 952
 Db 944 ATGCCAGGCTCAAGTATTACAAAGAGCTAAGCAAAAGCATTAATCCATACCTAGCGC 1003
 QY 953 AACGATTTGGAAGATATGCAATGATTAATTAATTTGATCTCTACCATCTGCTGTAAT 1012
 Db 1004 AACGATTTGGAAGATATGCAATGATTAATTAATTTGATCTCTACCATCTGCTGTAAT 1063
 QY 1013 TTATATCTCTCTTGAAGAAATTCATCCACAAGATTCATCAGGAAACCTTACATG 1072
 Db 1064 TTATATCTCTCTTGAAGAAATTCATCCACAAGATTCATCAGGAAACCTTACATG 1123
 QY 1073 CTTTCTGCTGCTGCTGTAACCTGTGTTCTGTGGACATCTAGGAAACATAGAAAT 1132
 Db 1124 CATTCTGCTGCTGCTGTAACCTGTGTTCTGTGGACATCTAGGAAACATAGAAAT 1183
 QY 1133 ATATTCAGCAGATTCATTCATCATGTGGAAGATTTAACTCTCCAGTAGAGTACCA 1192
 Db 1184 ATATTCAGCAGTTCATTCATCATGTGGAAGATTTAACTCTCCAGTAGAGTACCA 1243
 QY 1193 AGTATCTGAAGGAAGTGCACAAACAAATAGTTATACCTTAGTACTTTAACTGGAGGA 1252
 Db 1244 AATATCTGAAGGAAGTGCACAAACAAATAGTTATATCTTAGTACTTTAACTGGAGGA 1303
 QY 1253 AGGATTCAGTGAATCTTCCACAGATTTTGGAAATCAGATGTTGGCTTGCAATC 1312
 Db 1304 AGGATTTACTGTAACCTCCACAGGTTTGGGAATCAGATGATGCTGCGATGCGATC 1363
 QY 1313 ATGTGAAAAGCATCAAGATTAAGTCTGTGGTAATTTAGAGAAATGGTTTGGAAAT 1372
 Db 1364 ATGTGAAAAGCATCAAGATTAAGTCTGTGGCAATTTAGAGAAATGGTTTGGAAAT 1423
 QY 1373 AAAATTTTCATCACTGACACTGTGTAATATTTGATGAGATATCAACCAAGATTG 1432
 Db 1424 AAAATGT-----CATCATCTGTCACAGTAATGATTAATGAGACCAATTCAGTGTTC 1479
 QY 1433 AGGATGAAGAGATGCAATACATCTACTTGTGTGTCACAAATTTATTTTACACCTCTTC 1492
 Db 1480 AGGATGAAGAG-----ACACGCTGCAATTTGGTGCACATTAATTTCTCTGCTGCC 1534
 QY 1493 TAGGTAAGCTGATATTTTGGTGGAGATTTTAAAGCTCAGATGAGCAATCATTTCCA 1552
 Db 1535 TTGAGGACCATGATATTTTGAATTTTAAAGACCAAGATGAGCAAAATCATCTCA 1594

```

QY 1553 TTGGCTTTAAATTAATCCGTATATACCTTAATTAATGCACTGGACAGTAATTTATTTCTT 1612
DB 1595 TTGGCTTTAAATTAATCCGTATATATGATATGATACCAAGCACTGGA-TAATATTTATTCAT 1653
QY 1613 CATTATCATTTTGAACATTCGCTTTTTCACATTTTGTACTGTCTCAATTAATGAAGCTTG 1672
DB 1654 CATTTCATTTTGAACATTCGCTTTTTCACATTTTATTAATTAATGATGCTTAAGTAAGCTTG 1713
QY 1673 TGGTTGATTAATGTTTCCACACTGATGACCTGTTTAATCATTTTGGGAATGAAGATGC 1732
DB 1714 TGAATTTTAAATTTTCCACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1770
QY 1733 ACATCTTAAGTATGA-AAAATTTTCACTAGTATTAATTAATGCTAGTTCACACTTTGCA 1791
DB 1771 ATATCTTAATTAATTAAGAGCACTCTTTCACAAAGTATTAATTAATTAATTAATTAATTA 1830
QY 1792 TACATTAACAGAGAGAGAA--CATGTTGCGATTAATTAATTAATTAATTAATTAATTAATTA 1849
DB 1831 TAATGCTTAACAAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1890
QY 1850 TGAATGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1909
DB 1891 TGAACACTACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1936
QY 1910 TGAAGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1969
DB 1937 TGAATGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1988
QY 1970 GGAAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2029
DB 1989 TGAATGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2041
QY 2030 GGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2083
DB 2042 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2095

RESULT 6
AF230460 1705 bp mRNA linear ROD 12-JUL-2000
LOCUS Cricetulus griseus alpha(1.3)fucosyltransferase type IX (Fut9)
DEFINITION mRNA, complete cds.
ACCESSION AF230460.1 GI:9049663
VERSION AF230460.1
KEYWORDS Chinese hamster.
SOURCE Cricetulus griseus
ORGANISM Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus;
1 (bases 1 to 1705)
Patchak,S.K., Zhang,A., Shi,S. and Stanley,P.
Alpha(1.3)fucosyltransferases expressed by the gain-of-function
Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30
Arch. Biochem. Biophys. 375 (2), 322-332 (2000)
JOURNAL MEDLINE 2016953
PUBMED 10700388
REFERENCE 2 (bases 1 to 1705)
AUTHORS Patchak,S.K., Shi,S. and Stanley,P.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of
Medicine, Chanin 516, 1300 Morris Park Avenue, New York, NY 10461,
USA
FEATURES
source Location/Qualifiers
1..1705
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
1..1705
/gene="Fut9"
/note="c9Fut9"
62..1141
/gene="Fut9"
CDS

```

```

/Note="Fuc-IX"
/codon_start=1
/product="alpha(1.3)fucosyltransferase type IX"
/protein_id="AA082412.1"
/db_xref="GI:9049664"
/translation="MTSTSKGIIIRPEFLIVCVILACFMACLIILIKPNSWVSPMESA
SSVLKMKFEESTKIDYFNEETILVWVWPFQGFDTLSCQAMFNIOGCHLTDRLYNK
SHAVLILHRDISMDLNLPOARPEPFQKIMMNLSPYHTPOKSGIEHLFNLTLVRY
DSDIQVPGFLVSTNPPEVEYSEKELVWVSWMNPENHARVKNVLSKSEIHTY
GOAFGEYVNEKLIPTISTCKYLSFENSIRHDIYTEKLYNAFLAGSVVYLGSRN
YENYIPADSFIVEDYNSPELAKYLFKVDNNKNIYLSFYMRDFTVNPREFMESH
CIACDHVRHOREKSGNLEKWFNN"
BASE COUNT 515 a 336 c 320 g 534 t
ORIGIN
Query Match 42.9% Score 1147.6; DB 10; Length 1705;
Best Local Similarity 83.4% Pred No 1e-221;
Matches 1442; Conservative 0; Mismatches 259; Indels 29; Gaps 11;
QY 231 ACAGTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 290
DB 2 ACAGTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 57
QY 291 AATTATGACATCAACATCCCAAGAAATTCCTGCGCATTTTAATTTGTCGATTATCCT 350
DB 58 AATTATGACATCAACATCCCAAGAAATTCCTGCGCATTTTAATTTGTCGATTATCCT 117
QY 351 GGGCTGTTTCATGGCATGCTCTTCATTTTACATCAAACTTACCAACAGCTGGATCTTCAG 410
DB 118 GGGCTGTTTCATGGCATGCTCTTCATTTTACATCAAACTTACCAACAGCTGGATCTTCAG 177
QY 411 TCCATGAGATGAGCAGCTGCTGTGCTGAATAATGAATAAATCTTTTCCACCAAACTGA 470
DB 178 TCCATGAGATGAGCAGCTGCTGTGCTGAATAATGAATAAATCTTTTCCACCAAACTGA 237
QY 471 TTAATTTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 530
DB 238 TTAATTTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 297
QY 531 TACATCCGCGCAAGCAATGTTTCACATCCCAAGAGATGAGCATTCACAAAGGAGACCTTCACT 590
DB 298 TACATCCGCGCAAGCAATGTTTCACATCCCAAGAGATGAGCATTCACAAAGGAGACCTTCACT 357
QY 591 GTACACAAATCCCATGCGCTTCTGATCCATACCGAGACATCACTGGGATCTGACAAA 650
DB 358 GTACACAAATCCCATGCGCTTCTGATCCATACCGAGACATCACTGGGATCTGACAAA 417
QY 651 TTTTACTGACGAAGCTAGGCGCACCTTCAGAAATGATTTGGATGAATTTGGAATCAAC 710
DB 418 GTTACTGACGAAGCTAGGCGCGCCCTTTCAGAAATGATTTGGATGAATTTGGAATCAAC 477
QY 711 AACTCAGACTCCCAAAAGAGTGGCATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 770
DB 478 CACTCATTAACCCCAAAAGAGTGGCATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 537
QY 771 CCGTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 830
DB 538 CCGTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 597
QY 831 GTTTGAAGGCGCAAGCAAGAAATTTGTTGCTGGGCTGGTGGATTAATTAATTAATTAATTAATTAATTA 890
DB 598 GTTTGAAGGCGCAAGCAAGAAATTTGTTGCTGGGCTGGTGGATTAATTAATTAATTAATTAATTAATTA 657
QY 891 GGATGCGAGGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 950
DB 658 GGATGCGAGGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 717
QY 951 GGAACATTTGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1010
DB 718 GGAACATTTGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 777
QY 1011 AATTTATCTTCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1070

```


Db 778 ATTTATCTTCTTTGAAAACATTCACAAAGATTACATACCGAAAAGCTCTACAA 837
 QY 1071 TCGTTTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1130
 Db 838 TGGATTTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 897
 QY 1131 TTTATTTCCAGCAGATTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1190
 Db 898 TTTATTTCCAGCAGATTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 957
 QY 1191 AAGTATCTGAAAGGATGCAAAAACATTAAGTTATACCTAGTATACCTAGTATACCT 1250
 Db 958 AATATTTCTTAAGGAAGTGTACAAAACATTAAGTTATACCTAGTATACCTAGTAT 1017
 QY 1251 GAAGATTTTCACTGTAAATCTTCCAGATTTTGGGAATCATGATCATGATGATGAT 1310
 Db 1018 GAAGATTTTCACTGTAAATCTTCCAGATTTTGGGAATCATGATCATGATGATGAT 1077
 QY 1311 TCATGTGAAAAGGATCAAGATATTAAGTCTGTGTTGTTAGAGAAATGTTTGGAA 1370
 Db 1078 TCAGTAAAAAGGATCAAGATATTAAGTCTGTGTTGTTAGAGAAATGTTTGGAA 1137
 QY 1371 TTTAAATTTTTCATCTGACATGACATGATTAATTTTGTATGAGATATATCATCAAG 1430
 Db 1138 TTTATGTCTTCCATCTTCTCAATTTTAAAGAAATGATTAATTAATTAATTAATTA 1197
 QY 1431 TGAGATTAAGAAAGATGCAACATCTACTTCTTGTGCAATTTTATTTTATTCACCTC 1490
 Db 1198 TAAGATTAAGAAAGATGCAACATCTACTTCTTGTGCAATTTTATTTTATTCACCTC 1254
 QY 1491 TCTAGGATTAAGATGATTAATTTTGTGAGAA-TTTTTAAAGCTCAGCATGAGCAATCT 1549
 Db 1255 CTTTGAATTAAGATGATTAATTTTGTGAGAAATTTTAAAGCTCAGCATGAGCAATCT 1314
 QY 1550 CATTGCGTTTAAATTAATCTGTATTAATTAATTAATTAATTAATTAATTAATTA 1609
 Db 1315 TCAATTTGG-ATTAATTAATCTGTATTAATTAATTAATTAATTAATTAATTAAT 1372
 QY 1610 CTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1669
 Db 1373 CTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1431
 QY 1670 TTTGTTGTTGATTAATTTTGTGAGAAATTTTGTGAGAAATTTTGTGAGAAAT 1729
 Db 1432 TTTATGATTTT-CAATTTGTTTCTTCAATCAATCAATCAATCAATCAATCAAT 1488
 QY 1730 TGCACATCTTAAGTATGAAGAAATTTTCACTAGTATTAAGTATTAAGTATTAAG 1789
 Db 1489 TGGGTGCTTAAATTAAGAAAGCTTTTGTGCAAGTATTAATTAATTAATTAATTA 1548
 QY 1790 CATACCTTAATTAAGTATGAAGAAATTTTCACTAGTATTAAGTATTAAGTATTAAG 1847
 Db 1549 CATACCTTAATTAAGTATGAAGAAATTTTCACTAGTATTAAGTATTAAGTATTAAG 1608
 QY 1848 GATGATGAAGTATTAATTAATTTTGTGAGAAATTTTGTGAGAAATTTTGTGAGAA 1907
 Db 1609 GTTGA-----CAGTGTGATTAATTAATTTTGTGAGAAATTTTGTGAGAAAT 1655
 QY 1908 TGTGAAGTGAATTAATTAATTTTGTGAGAAATTTTGTGAGAAATTTTGTGAGAA 1957
 Db 1656 TGTGAAGTGAATTAATTAATTTTGTGAGAAATTTTGTGAGAAATTTTGTGAGAA 1705
 RESULT 7
 AF345993 1128 bp mRNA linear ROD 07-MAR-2001
 LOCUS Rattus norvegicus alpha1,3-fucosyltransferase IX (Fut9) mRNA,
 DEFINITION complete cds.
 ACCESSION AF345993
 VERSION AF345993.1 GI:13242183
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 AUTHORS 1 (bases 1 to 1128)
 TITLE Baboval,T., Henion,T., Kinnally,E. and Smith,F.I.
 JOURNAL Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX)
 PUBMED and comparison of the expression of fuc-TIV and fuc-TIX genes
 during rat postnatal cerebellum development
 J. Neurosci. Res. 62 (2), 206-215 (2000)
 11020213
 REFERENCE
 AUTHORS 2 (bases 1 to 1128)
 TITLE Smith,F.I. and Baboval,T.
 JOURNAL Direct Submission
 SUBMITTED (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver
 Center, 200 Trapeze Rd., Waltham, MA, USA
 FEATURES
 source
 1..1128
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 1..1128
 /gene="Fut9"
 49..1128
 /gene="Fut9"
 /codon_start=1
 /product="alpha1,3-fucosyltransferase IX"
 /protein_id="AAK16591.1"
 /db_xref="GI:13242184"
 /translation="MTSTSKGLRPLVLCIILGCVACILYIKPTNSWVSPMESA
 SSVLKKNFESKRTDFNEFTLLVWVPGQTFDDTSCAMRNIDGCHLNTLSLYNK
 SHAIVLIHHRDIDSLMDLNLPOQARPPOKIMWNLESPTTPKSGEHLNLTLYRR
 DSDIIVYGEFLVSTSPVEVEVPSKELVCMVSNVSNVSNVSNVSNVSNVSNVSNVSNV
 GOAFGEVYNDKMLIPTISTCKEYLSFENSIHDOYTERKYNAPLGSVYVVGPSRN
 YENYIPADSEFIHVEDNSPSLAKYIKEDVKNKKLXLYSFNMRKQFTVLLPFWESHA
 CLADPHVHKHOKYSVGNLEKFMN"
 BASE COUNT 324 a 256 c 228 g 320 t
 ORIGIN
 Query Match 35.8%; Score 958.2; DB 10; Length 1128;
 Best local Similarity 91.2%; Pred. No. 1.7e-183;
 Matches 1017; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 Db 260 TGCCTCCATGATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 319
 QY 14 TCGTTTCCATGATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 73
 Db 320 TTTGCCCATTTTAATTTGTGTGATTAATTTCTGGGCTTTTCAAGGATGCTTCTGATTT 379
 QY 74 TTTGCCCATTTTAATTTGTGTGATTAATTTCTGGGCTTTTCAAGGATGCTTCTGATTT 133
 Db 380 ACATCAAACTTACCAACAGCTGATCTTCACTCAATGAATGAATGAGCAGCAGCTGTGCTGA 439
 QY 134 ACATCAAGGCCACCAACAGCTGCTTCACTCAATGAATGAATGAGCAGCAGCTGTGCTGA 193
 Db 440 AAATGAAAACCTCTTTTCCACAAAACATGATTAATTAATTAATTAATTAATTAAT 499
 QY 194 AAATGAAAACCTCTTTTCCACAAAACATGATTAATTAATTAATTAATTAATTAATTA 253
 Db 500 GGGGTGGCCATTTGGGCAACCTTTGACCTTCACTTCACTTCACTTCACTTCACTT 559
 QY 254 GGGGTGGCCATTTGGGCAACCTTTGACCTTCACTTCACTTCACTTCACTTCACTT 313
 Db 560 AAGGATGCATCTCACAAAGGACGTTCACTGTACAAACAATGCCATGCGAGTCTGATTC 619
 QY 314 AAGGATGCATCTCACAAAGGACGTTCACTGTACAAACAATGCCATGCGAGTCTGATTC 373
 Db 620 ATCAGGAGACATCACTTGGGATCTGCAAAATTTTACCTGAGCAAGTAGGCAACCTTCC 679
 QY 374 ACCACAGACATCACTGAGGATCTGCAAAATTTTACCTGAGCAAGTAGGCAACCTTCC 433
 Db 680 AGAATGATTTTGGATTAATTTGGAATCACCACATCACTTCCCAAAAGAGTGCAATG 739
 QY 434 AGAATGATTTTGGATTAATTTGGAATCACCACATCACTTCCCAAAAGAGTGCAATG 493

QY	740	AGCACTTTGTTTAACTGACTGACTGACTTACCGCGCGGATGCAATATTCACAGGCGCTTATG	799
Db	494	AGCACTTTGTTTAACTGACTGACTTAAATTCGCGGATGCAATATTCACAGGCGCTTATG	553
QY	800	GCTTCTTCAGCGTATGACGCAAAATCCCTTCGTTGAGTGGCAGGCAAGAAAGAAATTCG	859
Db	554	GCTTCTTCAGCGTATGACGCAAAATCCCTTCGTTGAGTGGCAGGCAAGAAAGAAATTCG	613
QY	860	TGTGCTGGGTTGTGAGTAACTGGAACCTTGACATGCCAGAGTCAAGTATTCAGATGAGC	919
Db	614	TGTGCTGGGTTGTGAGTAACTGGAATCCCGACATGCCAGGAGTCAAGTATTCAGATGAGC	673
QY	920	TAAAGCAAAAGCATTTGAATTCATACCTTACGGGCAAGCATTTGGAGATATGTCAATGATA	979
Db	674	TCAAGAAAGATATTTGAATTCATACCTTATGGGCAAGCATTTGGAGATATGTCAATGATA	733
QY	980	AAAAATGATTCCTTACCATATCTGCTTGAATTTTATCTTCTTTGAAATTCATATCC	1039
Db	734	AAAAATGATTCCTTACCATATCTGCTTGAATTTTATCTTCTTTGAAATTCATATCC	793
QY	1040	ACAAGCATTTACATCAGCAAGAAAGCATATCAATAGCTTTTCTGCTGCTGCTGTACCGCTTG	1099
Db	794	ACAAGCATTTACATCAGCAAGAAAGCATATCAATAGCTTTTCTGCTGCTGCTGTACCGCTTG	853
QY	1100	TTCTGGGACCATCTGAGGGAACCTTGTGAAATATATTCACAGATTCATATTCATG	1159
Db	854	TTCTGGGACCATCTGAGGGAACCTTGTGAAATATATTCACAGATTCATATTCATG	913
QY	1160	TGGAGATTTATTAACCTCCGAGTGAAGCAATATCTGAAGCAAGTCCGACCAAAACA	1219
Db	914	TGGAGATTTATTAACCTCCGAGTGAAGTGGCAAAATATCTGAAGCAAGTCCGACCAAAACA	973
QY	1220	ATAAGTATACCTTAAGTACTTTAACTGAGGAGAGGATTTCACTGAAATCTCCACGAT	1279
Db	974	ATAAGTATATCTTAAGTACTTTAACTGAGGAGAGGATTTCACTGAAATCTCCACGAT	1033
QY	1280	TTTGGGAATCACATCATGTTTGGCTTGGCAATGTGAAAAGGCAATCAAGATATTAAGT	1339
Db	1034	TTTGGGAATCACATCATGTTTGGCAATGTGAAAAGGCAATCAAGATATTAAGT	1093
QY	1340	CTGTGGTAAATTTAGAAAATGTTTGGAAATTA	1374
Db	1094	CTGTGGCAATTTAGAAAATGTTTGGAAATTA	1128
RESULT 8			
LOCUS	AB035906	1707 bp	DNA linear
DEFINITION	Gallus gallus gene for cfrt9, complete cds.		VRT 05-JAN-2002
ACCESSION	AB035906		
VERSION	AB035906.1	GI:18146865	
KEYWORDS	cfrt9.		
SOURCE	Gallus gallus brain DNA.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archoosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
REFERENCE	1 (sites)		
AUTHORS	Kaneko,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.		
TITLE	The evolutionary history of glycosyltransferase genes		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1707)		
AUTHORS	Kaneko,M., Saitou,N. and Kitano,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-DEC-1999) Mka Kaneko, National Institute of Genetics, Laboratory of Evolutionary Genetics, Yata 1111, Mishima, Shizuoka, 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp, Tel:81-55-81-6790, Fax:81-55-81-6789)		
COMMENT	Sequence updated (29-Feb-2000).		
FEATURES	Location/Qualifiers		
SOURCE	1..1707		
	/organism="Gallus gallus"		

Query Match	Similarity	Score	DB	Length
Best Local	76.7%	Pred. No. 1,3e-155		
Matches 1061	Conservative 1	Mismatches 302	Indels 20	Gaps 4
gene				
CDS				
ORIGIN				
BASE COUNT	517 a	338 c	329 g	521 t
2 others				
Query Match	30.7%	Score 820.4	DB 5	Length 1707
Best Local	76.7%	Pred. No. 1,3e-155		
Matches 1061	Conservative 1	Mismatches 302	Indels 20	Gaps 4
gene				
CDS				
ORIGIN				
BASE COUNT	517 a	338 c	329 g	521 t
2 others				
Query Match	30.7%	Score 820.4	DB 5	Length 1707
Best Local	76.7%	Pred. No. 1,3e-155		
Matches 1061	Conservative 1	Mismatches 302	Indels 20	Gaps 4
gene				
CDS				
ORIGIN				
BASE COUNT	517 a	338 c	329 g	521 t
2 others				

```

Db 981 CAATTCCTACCTTCATTTGAAATTCATCCAAAGATTACATTCAGAGAACTTGA 1040
QY 1068 CAATGCTTTTGGCTGGCTCTGACCTGTTGTTGGACACATCAGGAGAACTATGA 1127
Db 1041 CATGCTCTTGGCTGGATCAGTACCAAGTTGACTGGGCCCTCCAGAGAAACATGCA 1100
QY 1128 GAATTAATTCAGCAGATTCATTCATCATGTGGAAGATTATTAATCTCCAGTAGCT 1187
Db 1101 GAATTCATTCAGCAGATTCATTCATCATGTGGAAGATTATTAATCTCCAGTAGCT 1160
QY 1188 AGCAAGTATTCAGAGAGAGTGCACAAAACATTAATGTTTACCTTAGTACTTAACTG 1247
Db 1161 TGCAGAAATCTCTGATGCTGACAAAATTAACAAGATGACCTAGTACTTCAACTG 1220
QY 1248 GAGGAGAGATTCACGTAATCTCCAGATTTGGGAAATCAGATGATGTTGGCTTG 1307
Db 1221 GAGGAGAGATTTCTGATGCTGATCTCTTAATCTGGGAATCTCATGATGCTTGGCTG 1280
QY 1308 CGATCATGTGAAAAGCATCAGAAATATAAGTCTGTGTAATTAAGAAATGCTTTTG 1367
Db 1281 TGATCATGTGAAAAGCATCAGAAATATAAGTCTGTGTAATTAAGAAATGCTTTTG 1340
QY 1368 GAATTA--AATTTTCATCCTTGCACACTGATTAATTTTGTGATGATATCATCCA 1425
Db 1341 GAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1400
QY 1426 AGTATTGAGATTAAGAA--GAGATGACATCACTACTACTTTTGTGTC 1469
Db 1401 AGTGGAGAGTAAGAAAGTAAGAAAAAGAGAGATGACACTCAGAACCTTAAGT 1460
QY 1470 CAATTAATTTTATCACCTCTCTAGGTTACGTATATTTTGGTGGAGATTTTAAAA 1529
Db 1461 CAAGTTTATCACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1520
QY 1530 GCGACAGATGAGAAATCAT--TGCATCGGTTTAAATTAATTCCTGTAATACCTAATTA 1588
Db 1521 AATTAATGAGCAGATGATCAAAACATGATTAATTAATTAATTAATTAATTAATTAAT 1579
QY 1589 TGCATGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1648
Db 1580 TGCATGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1639
QY 1649 GTAG 1652
Db 1640 GTGG 1643

RESULT 9
AB035905 1080 bp DNA linear VRT 05-JAN-2002
LOCUS xenopus laevis gene for xFUT9, complete cds.
DEFINITION AB035905
ACCESSION AB035905
VERSION AB035905.1 GI:18146863
KEYWORDS xFUT9.
SOURCE Xenopus laevis DNA.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (stiles)
AUTHORS Kaneko,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1080)
AUTHORS Kaneko,M., Saitou,N. and Kitano,T.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1999) Miki Kaneko, National Institute of
Genetics, Laboratory of Evolutionary Genetics, Yata 111, Mishima,
Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp,
Tel:81-559-81-6790, Fax:81-559-81-6789)
COMMENT
FEATURES
Sequence updated (29-Feb-2000).
Location/Qualifiers

```

```

source 1. 1080
/organism="Xenopus laevis"
/db_xref="taxon:8355"
gene 1. 1080
/gene="xFuc-TIX"
CDS 1. 1080
/gene="xFuc-TIX"
/codon_start=1
/product="xFUT9"
/protein_id="BAB82488.1"
/db_xref="GI:18146864"
/translation="MTSASKGLRFLAVSVGLTIPMLCLLYVPTNNWISSPIESA
HSVLKRNFESSKQNYNETVVLIVWPGQFELKSCOTLENIIGCHLTDRALYNK
SHAVLHHRDIDSMIDLINLEOQLRPFQKIMWNLSPTHTPKSGIIEHLFNLTYRR
DSDIQVYGFVSTKPEPDEVSKDKFVCWVSNMNPENHARVKYENELIKYIELITY
GOAFGYSLSDKSLPLTIAICKFYSFENSIRHDDYITERKYNALLNGSVPIYGPREN
YBNYIPADSFIVHEDPLSPRELADHLIMLNKRTDQYLSYNNRKHPTVNMHPFWSHA
CIACDHVHRHQRKYSVGNLEKRFNN"
BASE COUNT 332 a 231 c 202 g 315 t
ORIGIN
Query Match 25.3%; Score 676.8; DB 5; Length 1080;
Best Local Similarity 76.7%; Pred. No. 1.2e-126;
Matches 828; Conservative 0; Mismatches 252; Indels 0; Gaps 0;
QY 295 ATGACATCAACATCCAAAGAAATCTCGCCATTTTAAATGCTGCAATTCCTGGGC 354
Db 1 ATGACATCAACATCCAAAGAAATCTAGCGGCAATTTTACCTGTCAGCGGTGTTGATC 60
QY 355 TGTTCATGCGATGCTCTTCATTTACATCAACCTACACAGCTGATCTTACGCCA 414
Db 61 TGTTCATGATGATTTACTGATATACGTCACCAACCAACAACTGATTCAGTCTT 120
QY 415 ATGGAATCAGCCAGCTGTGCTGTAAGTAAGAAACGTCCTTCCACCAAACTGATTA 474
Db 121 ATGGAATCAGCAGTACGCTCTTAATTAATGAATAATTTTCTCTCTTAACGAAATTA 180
QY 475 TTTAATGAACACTATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 534
Db 181 TACAAATGAACCGCTGCTGTAATTTGGGTGTCCTTTGGGAGACTTTTGAACGAA 240
QY 535 TCGTCCAGCAATGTTCAACATCCAAAGATGCCATCTCAACAGGACCGTTCAGTAC 594
Db 241 TCGTCCAGCAATGTTTAATTTATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 595 AACAAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 654
Db 301 AACAAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 655 CCTCAGCAGAGTGGCCACCCCTCCAGAAATGATTTGGATTTGGATTCACCAACT 714
Db 361 CCCCAACAGCTGAGGCGACCTTCCAGAAATGATTTGGATTTGGATTCACCAACT 420
QY 715 CACACTCCCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
Db 421 CATACGCCCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 775 GATTGATATTCAGAGGCTTATGCTTCTTACAGGTAAGCAAAATCCCTCGTGT 834
Db 481 GACTCAGATATTCAGAGGCTTATGCTTCTTACAGGTAAGCAAAATCCCTCGTGT 540
QY 835 GAATGCCAAGCAAAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
Db 541 GAGGTGCCAAGCAAAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 895 GCCAGATCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
Db 601 GCCAGATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
QY 955 GCATTTGGAGAAATATGCAATGAATAATTTGATTCATCAATCTGCTGCTGCTGCTGCT 1014
Db 661 GCTTTGGTGAATCTGAGCGATTAAGCTGCTCTTCAATTAATGCTGCTGCTGCTGCTGCT 720

```



```
* 29077 29176: gap of 100 bp
* 29177 30093: contig of 917 bp in length
* 30094 30193: gap of 100 bp
* 30194 31099: contig of 906 bp in length
* 31100 31199: gap of 100 bp
* 31200 32097: contig of 898 bp in length
* 32098 32197: gap of 100 bp
* 32198 33113: contig of 916 bp in length
* 33114 33213: gap of 100 bp
* 33214 34125: contig of 912 bp in length
* 34126 34225: gap of 100 bp
* 34226 35133: contig of 908 bp in length
* 35134 35233: gap of 100 bp
* 35234 36150: contig of 917 bp in length
* 36151 36250: gap of 100 bp
* 36251 37171: contig of 921 bp in length
* 37172 37271: gap of 100 bp
* 37272 38159: contig of 888 bp in length
* 38160 38259: gap of 100 bp
* 38260 39156: contig of 897 bp in length
* 39157 39256: gap of 100 bp
* 39257 40188: contig of 932 bp in length
* 40189 40288: gap of 100 bp
* 40289 41210: contig of 922 bp in length
* 41211 41310: gap of 100 bp
* 41311 42222: contig of 912 bp in length
* 42223 42322: gap of 100 bp
* 42323 43218: contig of 896 bp in length
* 43219 43318: gap of 100 bp
* 43319 44233: contig of 915 bp in length
* 44234 44333: gap of 100 bp
* 44334 45246: contig of 913 bp in length
* 45247 45346: gap of 100 bp
* 45347 46261: contig of 915 bp in length
* 46262 46361: gap of 100 bp
* 46362 47279: contig of 918 bp in length
* 47280 47379: gap of 100 bp
* 47380 48276: contig of 897 bp in length
* 48277 48376: gap of 100 bp
* 48377 49282: contig of 906 bp in length
* 49283 49382: gap of 100 bp
* 49383 50305: contig of 923 bp in length
* 50306 50405: gap of 100 bp
* 50406 51314: contig of 909 bp in length
* 51315 51414: gap of 100 bp
* 51415 52335: contig of 921 bp in length
* 52336 52435: gap of 100 bp
* 52436 53163: contig of 728 bp in length
* 53164 53263: gap of 100 bp
* 53264 54210: contig of 947 bp in length
* 54211 54310: gap of 100 bp
* 54311 55199: contig of 889 bp in length
* 55200 55299: gap of 100 bp
* 55300 56212: contig of 913 bp in length
* 56213 56312: gap of 100 bp
* 56313 57287: contig of 975 bp in length
* 57288 57387: gap of 100 bp
* 57388 58245: contig of 858 bp in length
* 58246 58345: gap of 100 bp
* 58346 59236: contig of 891 bp in length
* 59237 59336: gap of 100 bp
* 59337 60272: contig of 936 bp in length
* 60273 60372: gap of 100 bp
* 60373 61289: contig of 917 bp in length
* 61290 61389: gap of 100 bp
* 61390 62299: contig of 910 bp in length
* 62300 62399: gap of 100 bp
* 62400 63322: contig of 923 bp in length
* 63323 63422: gap of 100 bp
* 63423 64336: contig of 914 bp in length
* 64337 64436: gap of 100 bp
* 64437 65354: contig of 918 bp in length
* 65355 65454: gap of 100 bp
```

```
* 65455 66352: contig of 898 bp in length
* 66353 66452: gap of 100 bp
* 66453 67362: contig of 910 bp in length
* 67363 67462: gap of 100 bp
* 67463 68370: contig of 908 bp in length
* 68371 68470: gap of 100 bp
* 68471 69379: contig of 909 bp in length
* 69380 69479: gap of 100 bp

Query Match      23.9%; Score 639; DB 2; Length 71396;
Best Local Similarity 85.0%; Pred. No. 1.3e-118;
Matches 764; Conservative 0; Mismatches 122; Indels 13; Gaps 4;

QY 366 ATGCTCTTCATTTATCATCAAACTTACCAACAGCTGATCTTCAGTCCATGGAATGCAAGC 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65456 ATCTCTGCGCCCTGTGGATGCCACCAACAGCTGCTGTCTAGTCCATGAGATGCTGC 65515

QY 426 CAGCTCTGCTGTAATAATGAAAACTTTCTTCCACCAAACTGATTTATTTAATGAAC 485
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65516 AAGTTCTGTCTGAAAAATGAAAAATTTCTTCTCCACAAAAAAGTATTTTAACGAAAC 65575

QY 486 TACTATTCGTGTGGGTGGCCATTTGGCAGACCTTGGACCTTACATCCGCCAAGC 545
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65576 TACCATTCGTGTTGGGTATGCCATTTGGGCGACCTTTGACCTTACATCTGCCAAGC 65635

QY 546 AATGTTCAACATCCAAAGATGCCATCTCACAGCAAGCCGTTCACTGTACAACAAATCCCA 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65636 AATGTTCAATATCCAAAGGTCATCTCACAGCAAGCCGTTCACTGTGTGACAAATCCCA 65695

QY 606 TGCAGTTCTGATCTATCCAGCAACATCACTAGTGGATCTGACAAATTTTACCTCAGCAAGC 665
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65696 TGGGTCCTGATTCACCATGAGACATCACTAGTGGATCTGACAAATTTTACCTCAGCAAGC 65755

QY 666 TAGCCACCCCTTCAGAAATGATTTGGATGATTTGAAATTCACCAACTCAGCACTCCCA 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65756 CAGGCCACCCCTTCAGAAATGATTTGGATGATTTGAAATTCAGCACTCCCACTCAGCACTCCCA 65815

QY 726 AAAGATGCGATTTAGACACTTTGTTAACTGACTGTACTTACCGCGCGTATTCAGATAT 785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65816 AAAGATGCGATTTAGACACTTTGTTAACTGACTGTACTTACCGCGCGTATTCAGATAT 65875

QY 786 CCAAGTGCCTTATGCTTCTTTCAGGTAAGCAAAATCCCTTGTGTGTAAGTCCCAAG 845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65876 CCAAGTGCCTTATGCTTCTTTCAGGTAAGCAAAATCCCTTGTGTGTAAGTCCCAAG 65935

QY 846 CAAAGGAATTTGTTGCTGGGTGGTGGATGATCTGGAACCCGAGCATCCGAGTCCCA 905
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65936 CAAAGGAATTTGTTGCTGGGTGGTGGATGATCTGGAACCCGAGCATCCGAGTCCCA 65995

QY 906 GTATTACATGAGCTAAGCAAAAGCATTTGAATTCATATCAGGCAAGCATTTGCTAGA 965
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65996 GTATTACATGAGCTAAGCAAAAGCATTTGAATTCATATCAGGCAAGCATTTGCTAGA 66055

QY 966 ATATGTCATGATTAATAATTTGATTCCTTACCATATCTGCTGTGTAATTTATCTTCTT 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66056 ATATGTCATGATTAATAATTTGATTCCTTACCATATCTGCTGTGTAATTTATCTTCTT 66115

QY 1026 TGAATTTCAATCCACAAGGATTCATCAGGAAAGCTTATCAATGCTTTCTGCTGG 1085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66116 TGAATTTCAATCCACAAGGATTCATCAGGAAAGCTTATCAATGCTTTCTGCTGG 66166

QY 1086 CTCTGACCTGTTGTTCTGGACCATCTAGGAAACATAGAAATTAATTTCCACAGCA 1145
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66167 CTCTGACCTGTTGTTCTGGACCATCTAGGAAACATAGAAATTAATTTCCACAGCA 66226

QY 1146 TTCTATTCATTC -ATGTGCAAGTTTAACTCT-CCCACTGAGCTAGCAAGATCTGAA 1202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66227 TTCTATTCATTCATTCGTTGCAAGTTTAACTCTCCCACTGAGCTAGCAAGATCTGAA 66286

QY 1203 GGAAGTCGACAAAATAATTAAGTT-ATACCTTGTACTTAACTGAGGAAGATTC 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66287 GGAAGTCGACAAAATAATTAAGTTGTTGTTACTTAACTGAGGAAGATTC 66345
```

RESULT 11
AC100674/c 71396 bp DNA linear HMC 22-NOV-2001
LOCUS AC100674
DEFINITION Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.
AC100674
AC100674.1 GI:17048040
KEYWORDS HIG, HIGS_Phrase0.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Mus musculus, clone RP23-167K24
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 71396)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgallier, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Chapel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardina, S.,
Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Haeferd, A., Horton, L., Hulme, W., Iller, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
Lamarez, R., Landers, T., Lebeck, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Ribback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, N., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, S., Stojanovic, N.,
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, N.

TITLE
JOURNAL
COMMENT
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center Project name: L15932
Center Clone name: 167_K_24

NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

788: contig of 788 bp in length
788 888: gap of 100 bp
888 1702: contig of 814 bp in length
1703 1802: gap of 100 bp
1803 2642: contig of 840 bp in length
2643 2742: gap of 100 bp
2743 3637: contig of 895 bp in length
3638 3737: gap of 100 bp

3738 4635: contig of 898 bp in length
4636 4735: gap of 100 bp
4736 5657: contig of 922 bp in length
5658 5757: gap of 100 bp
5758 6692: contig of 935 bp in length
6693 6792: gap of 100 bp
6793 7703: contig of 911 bp in length
7704 7803: gap of 100 bp
7804 8742: contig of 939 bp in length
8743 8842: gap of 100 bp
8843 9787: contig of 945 bp in length
9788 9887: gap of 100 bp
9888 10819: contig of 932 bp in length
10820 10919: gap of 100 bp
10920 11808: contig of 889 bp in length
11809 11908: gap of 100 bp
11909 12930: contig of 922 bp in length
12931 13847: contig of 917 bp in length
13848 13947: gap of 100 bp
13948 14886: contig of 939 bp in length
14887 14986: gap of 100 bp
14987 15931: contig of 925 bp in length
15932 16011: gap of 100 bp
16012 16903: contig of 892 bp in length
16904 17003: gap of 100 bp
17004 17906: contig of 903 bp in length
17907 18006: gap of 100 bp
18007 18931: contig of 925 bp in length
18932 19031: gap of 100 bp
19032 19945: contig of 914 bp in length
19946 20045: gap of 100 bp
20046 20951: contig of 906 bp in length
20952 21051: gap of 100 bp
21052 21945: contig of 898 bp in length
21950 22049: gap of 100 bp
22050 22950: contig of 901 bp in length
22951 23050: gap of 100 bp
23051 23945: contig of 895 bp in length
23946 24045: gap of 100 bp
24046 24955: contig of 910 bp in length
24956 25055: gap of 100 bp
25056 25980: contig of 923 bp in length
25981 26080: gap of 100 bp
26081 27052: contig of 972 bp in length
27053 27152: gap of 100 bp
27153 28067: contig of 915 bp in length
28068 28167: gap of 100 bp
28168 29076: contig of 909 bp in length
29077 29176: gap of 100 bp
29177 30093: contig of 917 bp in length
30094 30193: gap of 100 bp
30194 31099: contig of 906 bp in length
31100 31199: gap of 100 bp
31200 32097: contig of 898 bp in length
32098 32197: gap of 100 bp
32198 33113: contig of 916 bp in length
33114 33213: gap of 100 bp
33214 34125: contig of 912 bp in length
34126 34225: gap of 100 bp
34226 35133: contig of 908 bp in length
35134 35233: gap of 100 bp
35234 36150: contig of 917 bp in length
36151 36250: gap of 100 bp
36251 37171: contig of 921 bp in length
37172 37271: gap of 100 bp
37272 38159: contig of 888 bp in length
38160 38259: gap of 100 bp
38260 39156: contig of 897 bp in length
39157 39256: gap of 100 bp
39257 40188: contig of 932 bp in length
40189 40288: gap of 100 bp
40289 41210: contig of 922 bp in length

```

* 41211 41310: gap of 100 bp
* 41311 42222: contig of 912 bp in length
* 42223 42322: gap of 100 bp
* 42323 43218: contig of 896 bp in length
* 43219 43318: gap of 100 bp
* 43319 44233: contig of 915 bp in length
* 44334 44333: gap of 100 bp
* 44334 45246: contig of 913 bp in length
* 45247 45346: gap of 100 bp
* 45347 46261: contig of 915 bp in length
* 46262 46361: gap of 100 bp
* 46362 47279: contig of 918 bp in length
* 47280 47379: gap of 100 bp
* 47380 48276: contig of 897 bp in length
* 48277 48376: gap of 100 bp
* 48377 49282: contig of 906 bp in length
* 49283 49382: gap of 100 bp
* 49383 50305: contig of 923 bp in length
* 50306 50405: gap of 100 bp
* 50406 51314: contig of 909 bp in length
* 51315 51414: gap of 100 bp
* 51415 52335: contig of 921 bp in length
* 52336 52435: gap of 100 bp
* 52436 53163: contig of 728 bp in length
* 53164 53283: gap of 100 bp
* 53264 54210: contig of 947 bp in length
* 54211 54310: gap of 100 bp
* 54311 55199: contig of 889 bp in length
* 55200 55299: gap of 100 bp
* 55300 56212: contig of 913 bp in length
* 56213 56312: gap of 100 bp
* 56313 57287: contig of 975 bp in length
* 57288 57387: gap of 100 bp
* 57388 58245: contig of 858 bp in length
* 58246 58345: gap of 100 bp
* 58346 59236: contig of 891 bp in length
* 59237 59336: gap of 100 bp
* 59337 60272: contig of 936 bp in length
* 60273 60372: gap of 100 bp
* 60373 61289: contig of 917 bp in length
* 61290 61389: gap of 100 bp
* 61390 62299: contig of 910 bp in length
* 62300 62399: gap of 100 bp
* 62400 63322: contig of 923 bp in length
* 63323 63422: gap of 100 bp
* 63423 64336: contig of 914 bp in length
* 64337 64436: gap of 100 bp
* 64437 65354: contig of 918 bp in length
* 65355 65454: gap of 100 bp
* 65455 66352: contig of 898 bp in length
* 66353 66452: gap of 100 bp
* 66453 67362: contig of 910 bp in length
* 67363 67462: gap of 100 bp
* 67463 68370: contig of 908 bp in length
* 68371 68470: gap of 100 bp
* 68471 69379: contig of 909 bp in length
* 69380 69479: gap of 100 bp

```

Query Match 12.8%; Score 343.4; DB 2; Length 71396;

Best Local Similarity 64.3%; Pred. No. 7.2e-59; Matches 633; Conservative 0; Mismatches 303; Indels 49; Gaps 12;

```

QY 1102 CTGGGACCATCTAGGAAACATATGAGATTAATTCACACAGATTCATTCATCATGTG 1161
| | | | |
DB 30210 CAGTGGCGAGACATCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 30151
QY 1162 GAAGATTATTAACCTCCAGTGAAGTATCTGAAGGAAGTCGACAAAACAAT 1221
| | | | |
DB 30150 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 30091
QY 1222 AAGTTATACCTAGTACTTAACTGAGGAGGATTCACCTGAATCTTCCAC-GATT 1280
| | | | |
DB 30090 AAGTGTACTGTTACTTAACTGAGGAGGATTTTACTGTAACTTACACGCGGTT 30031

```

```

QY 1281 TTGGGAATCATGCATGTTGGCTTGCAATCATGTGAAAGGCAATCAAGATATTAAGTC 1340
| | | | |
DB 30030 TTGGGAATCATGCATGCCTGGCATGCAGACATGTAAAGGCAATCAAGATATTAAGTC 29971
QY 1341 TGTGTGAATTTGAGAAAGGTTTGGAAATTAATAATTTTCACTTTCACACTTGAT 1400
| | | | |
DB 29970 TGTGTGAATTTGAGAAAGGTTTGGAAATTTAAAGTTCACATGTGTCACA---GTG 29915
QY 1401 AAATATTTGATGAGATATCATCCAAAGTATTTGAGATTAAGAGATGACAACTACTACT 1460
| | | | |
DB 29914 AAAAAGATTATAGACATATTAATTCAGTTTCAGATTAAGAGAGAGAGAGAGAGAGAGAG 29860
QY 1461 TTGTGTCAAAATTTATTTTATTCACCCCTCTAGGGAACGATATATTTGGTGGAGA 1520
| | | | |
DB 29859 TTGGGGCACCCTTTATTTATTTCCCTCCCTCCCTGAGAGCAGCATATATTTGGT-GAAA 29801
QY 1521 TTTTAAAGCTGACGATGAGCATCATCTTCATTCGGTTTAAATTTTCCCTGATATACC 1580
| | | | |
DB 29800 TTTTAAAGATGAGATTAATTTGACATCATCAATTTGGTTTAAATTTCCCTGATATATAG 29741
QY 1581 TAAATATGTCACATGAGAGATTAATTTCTCATATTCATTTGTAACATTTGCTTTTTC 1640
| | | | |
DB 29740 TGATTAAGACACCTGGA-AATTAATTTTGTCTACTCTCATTTTATTAACATTTGTTTTC 29682
QY 1641 ACAATTTTGTAGTTGTCATTAATGATGATGATGATGATGATGATGATGATGATGATG 1700
| | | | |
DB 29681 A---TTTATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 29628
QY 1701 AGCGTTTATCTATTGGGAAATGAAAGATGACATCTTAAATATGCA-AAAATTTTAC 1759
| | | | |
DB 29627 AGATCTTTAACTTAATTTGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 29568
QY 1760 TAAATATTAACATGCTAGTTCATTAATGATGATGATGATGATGATGATGATGATGATG 1817
| | | | |
DB 29567 CAAGTATTAATTAAGATATTAATTTCTAGCTTGTATATCTTACAAAGAGATTAATTAAT 29508
QY 1818 CGATGTAATCTTAACTCTTTGACCTCAATGATGATGATGATGATGATGATGATGATGAT 1877
| | | | |
DB 29507 CAGTAACTTAACTCTTTGACCTCAATGATGATGATGATGATGATGATGATGATGATGATG 29457
QY 1878 GATCATTTTATTTTACTGTTTATCATATTTGAGAGGATTAATTAATTAATTAATTAAT 1937
| | | | |
DB 29456 ----ATAGAGTTTCTTACTGTACCATTTGTCAGGAGATTAATTAATTAATTAATTA 29402
QY 1938 AAGAAAGATTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1997
| | | | |
DB 29401 AAGAAAGATTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 29356
QY 1998 CATTAGCTGATTTGACGACTATTTATTCATGATGATGATGATGATGATGATGATGATG 2057
| | | | |
DB 29355 TATTAACTGTTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 29296
QY 2058 TTAATATCCAAATTTGTTGATTT 2082
| | | | |
DB 29295 TCAGATTAACAAATTAATGATTT 29271

```

RESULT 12

AB023627 1567 bp DNA linear VRT 02-APR-1999

LOCUS AB023627 cds. danlo re10 zrf1 gene for alpha(1,3)fucosyltransferase, complete

DEFINITION AB023627

ACCESSION AB023627.1 GI:4587295

VERSION alpha(1,3)fucosyltransferase.

KEYWORDS Danlo re10 DNA.

SOURCE Danlo re10

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (sites)

AUTHORS Kageyama, N., Natsuka, S. and Hase, S.

TITLE	Molecular cloning and characterization of two zebrafish alpha(1,3)fucosyltransferase genes developmentally regulated in embryogenesis			
JOURNAL	J. Biochem. (1999) In press			
AUTHORS	2 (bases 1 to 1567)			
REFERENCE	Kageyama, N., Natsuka, S. and Hase, S.			
JOURNAL	Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034, Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, Tel: 81-6-6850-5381, Fax: 81-6-6850-5381)			
FEATURES	Location/Qualifiers			
source	1.1567			
gene	/organism="Danio rerio" /db_xref="taxon:7955" /gene="zfT1" /gene="zfT1" /gene="zfT1" /product="alpha(1,3)fucosyltransferase" /protein_id="BA076706.1" /db_xref="GI:4587296"			
CDS	/codon_start=1 /translation="MDKITLTPPSKAAOKVILATFMLEISFCIFPVYVNPNTPEKFP ADVHTNGSCETECLELDVLRKONKCTIKNASNPSTPPANAPATARKDODETILLI WYVPGARDLGVCGSOFNIGCHLDDSDMOKAHGVHHRHDSADIPORPRPFO KIMWMMSPNSHNSPLINDGFLNLSFSPKSDIPYVGRLLIDATDEOKNTIPKRD KIVCWIVSNFOEHRKRSOYVNELVKHIEKAYGAGHNNVNDNDYGNVSSCKFYLSF ENSHIRHYTEKLFNPLALCTVAVVIGPRYKNEERIPRPAFIHVDPTPKRLADH KSLDNEQYKQYEMNRKHFISMSSFFGLEHACRCDYLRKKNHLYALDLKQWENG"			
BASE COUNT	499 a	334 c	293 g	441 t
ORIGIN				
Query Match	10.4%	Score 279.6;	DB 5;	Length 1567;
Best Local Similarity	58.3%	Fred. No. 2.5e-46;		
Matches 533;	Conservative 0;	Mismatches 369;	Indels 12;	Gaps 2;
QY 477	TAATGAACACTACTTCTGTTGGTGGTGGCCATTTGGCGACGCTTTGACCTTACATC	536		
DB 461	TCAGGAACCACTCTGTTGATCTGGGATATCTTTTGGGACAGCTTTATCTGGAGT	520		
QY 537	CTGCCAAGCAATGTTCAACATCCAAAGATCCATCTCACAACGACCCCTTCACTGTACA	596		
DB 521	GTGTGTTTCCAGTTCACATCCACGGCTGCTCATTTAACGACGACGAGATATGATCCA	580		
QY 597	CAATCCCATGACGTTCTGATTCATCCAGACATCATGTTGGAGTCTGACAAATTAC	656		
DB 581	GAAAGCCGATGAGATGATGTTTACACACAGACATCTGACGCGG-----ATTGGC	631		
QY 657	TCAGCAAGCTAAGCCCTTCCAGAAATGATTTGGATGATTAATCCACCACTCA	716		
DB 632	GCAATCCCTCGACACACTTCCAAAGATGATGATGATGATGATGATGATGATGATG	691		
QY 717	CATCCCAAGCAAGTGGCATGAGCATTTGTTTAACTCTGACTCTGACTTACCCCGCTGA	776		
DB 692	TTCGCAATCCAAACAGCTTGAATTAATGATGATGATGATGATGATGATGATGATG	751		
QY 777	TTCGCAATCCAAAGTGCCTTA--TTGGCTTCTGACGATGACGACCAAAATCCCTTCTGT	833		
DB 752	TTCGCAATTCACAGTGCCTTACGTCGACATCATGATGACGACGACGACGACGACGAC	811		
QY 834	TGAAGTCCCAAGCAAGAAATTTGGTGTGGTTGGATGATGATGAAACCCCTGAGCA	893		
DB 812	CACATCCCAAGCAAGAAATTAACATGTTTCTGATGATGATGATGATGATGATGATG	871		
QY 894	TGCCAGATCAAGTATTCATGAGCTAGCAAAAGACATTTGAAATCCATACCTACGCGCA	953		
DB 872	CAAACTCCCAAGCAAGAAATTTGGTGTGGTTGGATGATGATGATGATGATGATGATG	931		
QY 954	AGCAATTCGAGAAATATGTCATGATTAATAATTTGATTCCTACCATATGCTTGTGAAT	1013		
DB 932	GCACCTCAATACCGGGTAAATGATGATGATGATGATGATGATGATGATGATGATG	991		

QY 1014	TTATCTTCTCTTGAATAATTCATCCAGAGATTACATCAGGAAAGCTTACATGC	1073
DB 992	CTACCTTCTCTTGAATAATTCATCCAGAGATTACATCAGGAAAGCTTACATGC	1051
QY 1074	TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1133
DB 1052	TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1111
QY 1134	TATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	1193
DB 1112	TATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	1171
QY 1194	GATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	1253
DB 1172	TCACCTTAAATCCTTGACCAAAACGAGACAGTACAGACATTCATTCATTCATTC	1231
QY 1254	GGATTCAGCTGTAATCTTCACAGATTTTGGATGATCATGATGATGATGATGATGAT	1313
DB 1232	ACATTTTATTCATGATGATTCATCATTCATTCATTCATTCATTCATTCATTCATTC	1291
QY 1314	TGTGAAGAAGCATCAAAATTAAGCTGTTGATTTAGAAATGCTTTGAAATTA	1373
DB 1292	TCTTAAAGGAATTAAGCATATTAAGCTGTTGATTTAGAAATGCTTTGAAATTA	1351
QY 1374	AAATTTTTCATGC	1387
DB 1352	ATGGGATTGTTGAC	1365
RESULT 13		
AB023628	1832 bp	DNA
LOCUS		
DEFINITION	Danio rerio zft2 gene for alpha(1,3)fucosyltransferase, complete cds.	
ACCESSION	AB023628	
VERSION	AB023628.1	
KEYWORDS	alpha(1,3)fucosyltransferase.	
SOURCE	Danio rerio DNA.	
ORGANISM	Danio rerio	
REFERENCE	1 (sites)	
AUTHORS	Kageyama, N., Natsuka, S. and Hase, S.	
TITLE	Molecular cloning and characterization of two zebrafish alpha(1,3)fucosyltransferase genes developmentally regulated in embryogenesis	
JOURNAL	J. Biochem. (1999) In press	
AUTHORS	2 (bases 1 to 1832)	
TITLE	Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0034, Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, Tel: 81-6-6850-5381, Fax: 81-6-6850-5381)	
FEATURES		
source	1.1832 Location/Qualifiers	
gene	/organism="Danio rerio" /db_xref="taxon:7955" /gene="zfT2" /gene="zfT2" /gene="zfT2" /product="alpha(1,3)fucosyltransferase" /protein_id="BA076707.1" /db_xref="GI:4587298" /translation="MDKITLTPPSKAAOKVILATFMLEISFCIFPVYVNPNTPEKFP ADVHTNGSCETECLELDVLRKONKCTIKNASNPSTPPANAPATARKDODETILLI WYVPGARDLGVCGSOFNIGCHLDDSDMOKAHGVHHRHDSADIPORPRPFO KIMWMMSPNSHNSPLINDGFLNLSFSPKSDIPYVGRLLIDATDEOKNTIPKRD KIVCWIVSNFOEHRKRSOYVNELVKHIEKAYGAGHNNVNDNDYGNVSSCKFYLSF ENSHIRHYTEKLFNPLALCTVAVVIGPRYKNEERIPRPAFIHVDPTPKRLADH KSLDNEQYKQYEMNRKHFISMSSFFGLEHACRCDYLRKKNHLYALDLKQWENG"	
CDS	332..1510	

Tue Oct 8 10:18:31 2002

us-09-744-748-5.rge

Page 20

D	b	436	TGCGATATCTTCATCCGCCCTACGGCTTCTT --- ATCTCCCAAAAGCCTGCCAAAATTAGTC	552
O	y	838	GTCGCACAAAGAAAGAAATTGGTGCTGGGTGTGTGAGTAAGTAGTGAACCCTGAGCATGCG	897
D	b	553	CTGCCGCCACAAGAAAGCTGGTGCCCTGGGTGATCAGCACMACGTGMAACGAGCACAGAA	612
O	y	898	AGAGTCAAGTATTACATAGACTAAGCAAAACATTTGAATCCATTAACCTACGAGGCAAGCA	957
D	b	613	AGGGTCAGTATTACACAGAGCTGAGAACATCGTCGAATATGACGTTTACGGCC --- GG	669
O	y	958	TTTGGAGAATATGTCATATATATAATTGATTCCTACCATATTCCTCTGTAAATTTTAT	1017
D	b	670	TATGGGCTGGATTTTAAAAAGAGCAACAATCGCAAGACCGCTCGAGATPAAATTTTAC	729
O	y	1018	CTTTCCTTTGAAAATTCATCCACAAGAGATTACATCAGGAAAGCTAT --- ACATGCT	1074
D	b	730	TTGGCTTTTGAATAATTCCTTGACACAGGATTACATCCAGAAAGCTGTGAGGAAACGGG	789
O	y	1075	TTTTGGCTGGCTCTGTACCTGTTGTCGGGACCATCTAGAGAAACATATGAGTATT	1134
D	b	790	TTTAAATCCAACGCCATCCCCTCATGTCATGGGGCCAGCCGCTATTAATCATGAGATGTC	849
O	y	1135	ATTCACAGAGATTCATATTCATTTGAGAAATTTAACTCTCCAGTAGTAGTAAAG	1194
D	b	850	ATACCGCGACGCTTTTATTACGTGAGCATTTCTCCAGCCCCAGAGAACTGGCCATG	909
O	y	1195	TATCTAGAGAAAGTGCACAAAACAAATAGTATTAACCTAGTACTTTTAACTAGAGAAAG	1254
D	b	910	TACCTGAAACACCTGGATTAATAAACATTCACCTGTACCGGAGGATATTACAGTGGAAAG	969
O	y	1255	GATTCACTGTAAATCTTCCACAGATTTTGGGAAATCAATGCATGTGTGGCTGCATCAT	1314
D	b	970	CGTACGACGTGCACAGTGCACGTCTTCTGGGACAGACANTTACTGCACCCGCTGCCAGAC	1029
O	y	1315	GTGAA 1319	
D	b	1030	GTCAA 1034	
R	E	S	RESULT 15	
L	O	C	AF288369	1264 bp DNA linear VMT 02-DEC-2001
A	C	C	Gallus gallus alpha (1..3) fucosyltransferase gene, complete cds.	
V	E	R	AF288369	
K	E	Y	AF288369.1 GI:17224489	
S	O	U	Chicken.	
O	R	G	Gallus gallus	
R	E	F	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
I	N	F	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;	
R	E	F	Phasianinae; Gallus.	
P	R	E	1 (bases 1 to 1264).	
A	D	H	Foley/A.C., Schmitz/B., Stern,C.D. and Strelt,A.	
J	O	U	Expression of fucosyltransferases in chick embryos	
R	E	F	Unpublished	
I	T	L	2 (bases 1 to 1264)	
R	E	F	Foley/A.C., Schmitz,B., Stern,C.D. and Strelt,A.	
A	D	H	Direct Submission	
J	O	U	Submitted (19-JUL-2000) Genetics and Development, Columbia	
R	E	F	University, 701 West 168th Street, New York, NY 10032, USA	
F	E	A	Location/Qualifiers	
E	A	T	1..1264	
S	O	U	/organism="Gallus gallus"	
C	D	S	<db_xref="taxon:9031"	
C	D	S	<1..1062	

```

/codon_start=1
/product="alpha (1,3) fucosyltransferase"
/protein_id="FA136994.1"
/db_xref="GI:17224490"
/translational_table="SWKKLLTFLIFLFSFTSCFCFASLRQFRPPEEDYNSTSLTERLE
AGGTSMLPKRSRLTLILMTWPGHNSRNCSELDPLDNGHLLPTTNKSMQAQADVIT
HHRWSPRIKRESLTIPALPSQRNIGWNRSSPHSLGMADHLFM/MSYRSDSIV

```

BASE COUNT 286 a 355 c 319 g 304 t
ORIGIN

Query Match	8.2%	Score 220.2	DB 5	Length 1264
Best Local Similarity	58.0%	Pred. No. 2.4e-34		
Matches 447	Conservative	0	Mismatches 318	Indels 6
			Gaps	
QY 487	ACTTTCCTGGGTGGGTGGTGGCCATTTTGGGCAACCTTTGACCTTTCATCTCAGCCCAAGCA	546		
Db 175	ACCATCTGGCTGGTGGACATGAGCCCTTGGCTGACCGCTTTAAATTTCAGCAACTGCTCAGAG	234		
QY 547	ATGTTCAAATCCAAAGATATCCCATCTCACAAGGACCGCTTCACTGTACAAACAATTCCTAT	606		
Db 235	CTCTTTGACACCCGGAATGCACTTTCACCAACCGACGCTGGTCCCAAAAGGAGAT	294		
QY 607	GCAATTCGTATCATCTACCGCAGACATCAGTTGGGATCTGACAAATTTTACCTCAGCAGCT	666		
Db 295	GCTGTGATTTATGCTATCCACAGGATGTGTATTTGGGACAGAGAGGGCTGTCCGGATCCCC	354		
QY 667	AGGGCAACCCCTTCAGAAATGGATTTGGATGATCAATTTTGGATCAACAATCAGCTCCCAAC	726		
Db 355	AGGCTCCCATCCAGCCCTGGATCTGGTTCACATGAGATCCCAAGTCAATTCCTCCAAAC	414		
QY 727	AAGAGTGGCATTTGAGCACTGTTTAACTGACACTCTACTTACCGCCGCTATTCAGATATC	786		
Db 415	TTAGATGCCATGACACCACTTCCTCAACCTGACTATGTCTGACCGGAGAGACTCGGACATC	474		
QY 787	CAATGCGCTTATGGCTCTCTGACGGTAAGACAAATCCCTGCTGTGTTTAACTGGCAAGC	846		
Db 475	TTCAACCCCTATGGGAGAGCTGACGTGCTCAGGAGACCCCAAGCCCTTCAGCATCCCAACC	534		
QY 847	AAAGAGAAATTTGGTGTGCTGGCTGTTGGATGAACTGGGAACCCCTGAGATCCGCAAGCTCAG	906		
Db 535	AAGACCAAGATGTTGGCTCTGGTGGTGTCACTGTAATTTGGAGACAAAGCTCCCAAGGGGTGAG	594		
QY 907	TATTACAAATGAGCTAAGCAAAABACATTTGAAATTCATACCTACGGGCGACAGCATTTGGAGAA	966		
Db 595	TACTACGAGAGGCTGAAGAAACACATCTCGTGGATGTGTACGGGGG-GCCAAACACTTGCC	653		
QY 967	TATCTCAATGATAAAAATTTGATTTCTTACCATATCTGCTGTAAATTTTATCTTTCCTTT	1026		
Db 654	CTCTTCATG--GACATGCTCCACAAACCGTATCCCAATACAACTTCTACTTGGCCCTTT	711		
QY 1027	GAAAATTCATATCCACAAGATTTACATCAGCGGAAAGAACTATAC---ATAGCTTTTGTGGCT	1083		
Db 712	GAGAACTCAAGAGCAAAAGACTACATCAGTAGAAGAACTCTGGAGGAATCCCTGTCTCTCT	771		
QY 1084	GGCTCTTACCTCTTGTCTTGGGAGACCATCTAGGAGAAACATAGAGAAATTAATTTCCAGAA	1144		
Db 772	GGCACTATCCCACTGTTGTTCTGGGCCACCTCGAGAAAACTATGAGAGACTTCTTTGCCCTCT	831		
QY 1144	GATTCATCTCATGTATGGGAAATTTTAACTCTCCCGAGSACTAGCAAAAGTATGTGAG	1203		
Db 832	GACTCCTTATCCACAGTCAATGACTTTGGCAGTGTCTGGTGAAGTGGCGAGTACTCTGAG	891		
QY 1204	GAAGTGCAGAAAAACAATAGTATATACCTTATGTTACTTAACTGGAGAG	1254		
Db 892	GAATGAGCCAGGAGACCGGAAGTACACAGCGCTACTCTCAATGGGCGAGG	942		

Search completed: October 6, 2002, 04:42:32
Job time : 3757.32 secs

Db 1501 CGGTATATTTTGGTGGAGATTTTAAAGCTCAGCATGACATCTTCCATTGCGTT 1560
 QY 1561 TAAATTAATCCGTATATACCTAAATATATGCTAGTGGAGATTAATTCCTCATTAATCA 1620
 Db 1561 TAAATTAATCCGTATATACCTAAATATATGCTAGTGGAGATTAATTCCTCATTAATCA 1620
 QY 1621 TTGTAAACATTCGCTTTTCCATTTTGTAGTTCCTCAATTAAGTGAAGCTTGGTGTGA 1680
 Db 1621 TTGTAAACATTCGCTTTTCCATTTTGTAGTTCCTCAATTAAGTGAAGCTTGGTGTGA 1680
 QY 1681 TTATTTGTTCCACATGATCAGCTGTTTAATCTATTTGGGAATGAAGATGACATCTTA 1740
 Db 1681 TTATTTGTTCCACATGATCAGCTGTTTAATCTATTTGGGAATGAAGATGACATCTTA 1740
 QY 1741 AAGTATGAAAAATTTTCTACTAGTATTAACATCTAGTTTCAACTTTGCTACTATTAAC 1800
 Db 1741 AAGTATGAAAAATTTTCTACTAGTATTAACATCTAGTTTCAACTTTGCTACTATTAAC 1800
 QY 1801 AGAGGAAGAATGTTGGCGATTTGAATTTCTAAGCTTTTGCCTTAAGATGAAGTGA 1860
 Db 1801 AGAGGAAGAATGTTGGCGATTTGAATTTCTAAGCTTTTGCCTTAAGATGAAGTGA 1860
 QY 1861 TATTAAGTCTATTTGATGATCTATTTTCTTACCTGTTTATCAATTTTGAAGGTGAAA 1920
 Db 1861 TATTAAGTCTATTTGATGATCTATTTTCTTACCTGTTTATCAATTTTGAAGGTGAAA 1920
 QY 1921 TTATTCATGAGATGAATTAAGAAAGATATGAAGAGAACTGTTCTATTCAGGAAGCTATTA 1980
 Db 1921 TTATTCATGAGATGAATTAAGAAAGATATGAAGAGAACTGTTCTATTCAGGAAGCTATTA 1980
 QY 1981 GACTTCTCATTTTTCATTAAGCTGATTTGCAAGCTACTTATTCATGCTGCTTAATTT 2040
 Db 1981 GACTTCTCATTTTTCATTAAGCTGATTTGCAAGCTACTTATTCATGCTGCTTAATTT 2040
 QY 2041 AAATTTATCAAGATTTTAAATATCAATTTGTTGATTTTCAGAGCACTGGGAAGTAA 2100
 Db 2041 AAATTTATCAAGATTTTAAATATCAATTTGTTGATTTTCAGAGCACTGGGAAGTAA 2100
 QY 2101 TCCCAATTAATCTTTTGAAGAAATCTAAGACAGTCTTCTGCTACTGATGACACTCATTTGT 2160
 Db 2101 TCCCAATTAATCTTTTGAAGAAATCTAAGACAGTCTTCTGCTACTGATGACACTCATTTGT 2160
 QY 2161 CATATAAATAAATAATATTTCTCCAAATTAACAAAGAAATGATACCTATTAATATATTT 2220
 Db 2161 CATATAAATAAATAATATTTCTCCAAATTAACAAAGAAATGATACCTATTAATATATTT 2220
 QY 2221 ATAAATGCTGATTAATGAACAATGTTTAATATGATCAATTTTAAGATTTTTTCTGA 2280
 Db 2221 ATAAATGCTGATTAATGAACAATGTTTAATATGATCAATTTTAAGATTTTTTCTGA 2280
 QY 2281 AGCCCTAATTAATTAATGAGCTTAAATTTACATATGATATGAATTTGGCTCATTAATG 2340
 Db 2281 AGCCCTAATTAATTAATGAGCTTAAATTTACATATGATATGAATTTGGCTCATTAATG 2340
 QY 2341 ATAGAGCCCTATCTTTGATTTGAGTCTATCATTTAAGAGAGCTTAATTAATATATCAT 2400
 Db 2341 ATAGAGCCCTATCTTTGATTTGAGTCTATCATTTAAGAGAGCTTAATTAATATATCAT 2400
 QY 2401 CAAGGATTAATTAATTAAGAGCTTAAATTAATTAAGTGGGATATATTAAGAAACACACAG 2460
 Db 2401 CAAGGATTAATTAATTAAGAGCTTAAATTAATTAAGTGGGATATATTAAGAAACACACAG 2460
 QY 2461 TGTATGACAGAGTAAAGATCTCAATGACATTTGTTGGATGAATTAATAATGCAATATGA 2520
 Db 2461 TGTATGACAGAGTAAAGATCTCAATGACATTTGTTGGATGAATTAATAATGCAATATGA 2520
 QY 2521 ATTCCCAAGAAAAATGATTTGTTCAAGAGAGTGCAGTTCATCTTTGAAGTACTAATATGG 2580
 Db 2521 ATTCCCAAGAAAAATGATTTGTTCAAGAGAGTGCAGTTCATCTTTGAAGTACTAATATGG 2580
 QY 2581 AGATGACTTTTATATCCATTTTGGTAAATTAATTCATATGACATATGACATATGATGTGT 2640
 Db 2581 AGATGACTTTTATATCCATTTTGGTAAATTAATTCATATGACATATGACATATGATGTGT 2640

QY 2641 TCAGGGCTTTATAGAACCAATAAACCCTACCAATTAC 2676
 Db 2641 TCAGGGCTTTATAGAACCAATAAACCCTACCAATTAC 2676
 RESULT 2
 AA292646
 ID AA292646 standard; cDNA; 2822 BP.
 AC AA292646;
 AC AA292646;
 DT 05-JUN-2000 (first entry)
 DE Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO.4.
 DE Human alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
 KW treatment; cancer; human; ss.
 OS Homo sapiens.
 FH Key location/Qualifiers
 FT CDS 1289..2368
 FT /*tag= a
 FT /product= "Human alpha-1,3-fucosyltransferase"
 PN WO200006708-A1.
 PD 10-FEB-2000.
 PD 29-JUL-1999; 99WO-JP04092.
 PE 29-JUL-1998; 98JP-0213823.
 PR (KYOM) KYOMA HAKKO KOGYO KK.
 PA Narimatsu H, Kudo T, Sasaki K;
 PI WPI: 2000-183120/16.
 DR P-FSDB: AAY80996.
 DR Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT epitope and an antibody recognizing it useful for diagnosis of brain
 PT and kidney disease and cancer.
 PS Claim 4; Page 134-142; 172pp; Japanese.
 PS The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognize alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3- fucosyltransferase. Sequences

CC AA292646 and AA292647 represent cDNAs encoding human
 CC alpha-1,3-fucosyltransferase (AA180996).
 XX Sequence 2822 BP; 918 A; 485 C; 477 G; 942 T; 0 other;

Query Match 57.6%; Score 1542; DB 21; Length 2822;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1542; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 GAAAAATATGACATCAACATCCAAAGAAATTCGCCCCATTTTAAATGTGTGATTA 346
 Db 1281 GAAAAATATGACATCAACATCCAAAGAAATTCGCCCCATTTTAAATGTGTGATTA 1340
 QY 347 TCCGCGCTGCTTTCATGAGCATGCTCTCTCATTTTACATCAACCTGACAGCATCT 406
 Db 1341 TCCGCGCTGCTTTCATGAGCATGCTCTCTCATTTTACATCAACCTGACAGCATCT 1400
 QY 407 TCAATGCCAATGSAATGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
 Db 1401 TCAATGCCAATGSAATGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1460
 QY 467 CTGATTTTAAATGAAATCAATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
 Db 1461 CTGATTTTAAATGAAATCAATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1520
 QY 527 ACCTTACATCTGCGAAGCAATGTTCAACATCCAAAGATGCTGCTGCTGCTGCTGCT 586
 Db 1521 ACCTTACATCTGCGAAGCAATGTTCAACATCCAAAGATGCTGCTGCTGCTGCTGCT 1580
 QY 587 CACTGTACAACAAATCCCATGAGTCTGATTCATCCAGACATCAATGCTGCTGCTGCT 646
 Db 1581 CACTGTACAACAAATCCCATGAGTCTGATTCATCCAGACATCAATGCTGCTGCTGCT 1640
 QY 647 CAAATTTACCTCAAGCAAGTAGGCAACCTTCCAGAAATGAAATGAAATTTGGAAT 706
 Db 1641 CAAATTTACCTCAAGCAAGTAGGCAACCTTCCAGAAATGAAATGAAATTTGGAAT 1700
 QY 707 CACCAACTCAACATCCCAAAAGATGAGCATGAGCATGCTGCTGCTGCTGCTGCTGCT 766
 Db 1701 CACCAACTCAACATCCCAAAAGATGAGCATGAGCATGCTGCTGCTGCTGCTGCTGCT 1760
 QY 767 ACCGCGCTGATTCAGATATCCAAAGTGCCTATGAGCTCTTGAAGGTAAGCAAAATCCCT 826
 Db 1761 ACCGCGCTGATTCAGATATCCAAAGTGCCTATGAGCTCTTGAAGGTAAGCAAAATCCCT 1820
 QY 827 TCGTGTGAGTGGCCAAAGCAAGAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 886
 Db 1821 TCGTGTGAGTGGCCAAAGCAAGAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1880
 QY 887 CTGAGCATGCGCAAGTCAATATTAACATGAGCTAAGCAAAAGCATTTGAATCCATACCT 946
 Db 1881 CTGAGCATGCGCAAGTCAATATTAACATGAGCTAAGCAAAAGCATTTGAATCCATACCT 1940
 QY 947 ACGGGCAAGCATTTGAGAAATATGTCATGATATAAAATTTGATTCCTACCATATCTGCTT 1006
 Db 1941 ACGGGCAAGCATTTGAGAAATATGTCATGATATAAAATTTGATTCCTACCATATCTGCTT 2000
 QY 1007 GTAATTTTATCTTCTTCTTGAATATCAATCCCAAGATTAACATCAACGAAAGCTAT 1066
 Db 2001 GTAATTTTATCTTCTTCTTGAATATCAATCCCAAGATTAACATCAACGAAAGCTAT 2060
 QY 1067 ACAATGCTTTTGGCGTGTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1126
 Db 2061 ACAATGCTTTTGGCGTGTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180
 QY 1127 AGAATTTATTTCCAGAGATTCATCATTCATGATGGAAGTATTAATCTCCAGAGTACG 1186
 Db 2121 AGAATTTATTTCCAGAGATTCATCATTCATGATGGAAGTATTAATCTCCAGAGTACG 2180
 QY 1187 TACCAAGTATCTGAAGGAAGTGCACAAAACAAATTAAGTTAACTTACTTTAACT 1246
 Db 2181 TACCAAGTATCTGAAGGAAGTGCACAAAACAAATTAAGTTAACTTACTTTAACT 2240

QY 1247 GGAGGAAGATTTTCACTGTAAATCTTCCACAGATTTTGGGAATCCATGATGCTTTGCTT 1306
 Db 2241 GGAGGAAGATTTTCACTGTAAATCTTCCACAGATTTTGGGAATCCATGATGCTTTGCTT 2300
 QY 1307 GCGATCATGTGAAAAGCAATCAAGAAATATAGTCTGTTGTAATTTAGAAATGCTTTT 1366
 Db 2301 GCGATCATGTGAAAAGCAATCAAGAAATATAGTCTGTTGTAATTTAGAAATGCTTTT 2360
 QY 1367 GGAATTTAAATTTTTCATCACTGTCACACTGTCATGATTAATTTTGAATGATATCCCA 1426
 Db 2361 GGAATTTAAATTTTTCATCACTGTCACACTGTCATGATTAATTTTGAATGATATCCCA 2420
 QY 1427 GTATTGAGATTAAGAGATGCAATCAATCTTTTGTGTCACATTTATTTATGAC 1486
 Db 2421 GTATTGAGATTAAGAGATGCAATCAATCTTTTGTGTCACATTTATTTATGAC 2480
 QY 1487 CCTCTGAGGTAAAGTATATTTTGTGAGATTTTAAAGCTCAGCATGAGCAATC 1546
 Db 2481 CCTCTGAGGTAAAGTATATTTTGTGAGATTTTAAAGCTCAGCATGAGCAATC 2540
 QY 1547 ATTCCATTTGGTTTAAATATTCCTGTATATACCTAATTAATGCACTGGAAGTAAAT 1606
 Db 2541 ATTCCATTTGGTTTAAATATTCCTGTATATACCTAATTAATGCACTGGAAGTAAAT 2600
 QY 1607 ATTCTCATTTATCATTTTGTAAACATTTGCTTTTTCACATTTTGTAGTGTCCATATGTA 1666
 Db 2601 ATTCTCATTTATCATTTTGTAAACATTTGCTTTTTCACATTTTGTAGTGTCCATATGTA 2660
 QY 1667 AGCTTGTGTTGATTAATTTTTCACACTGATCAAGCTGTTTAAATCTAATTTGGAAATGA 1726
 Db 2661 AGCTTGTGTTGATTAATTTTTCACACTGATCAAGCTGTTTAAATCTAATTTGGAAATGA 2720
 QY 1727 AGATGCAATCTTAAAGTATGAAATTTTCACTAATGATTAATGCTATGCTTCACT 1786
 Db 2721 AGATGCAATCTTAAAGTATGAAATTTTCACTAATGATTAATGCTATGCTTCACT 2780
 QY 1787 TTGCAATCTATTAACAGAGGAAGCAATGTTGGATGAAATTC 1828
 Db 2781 TTGCAATCTATTAACAGAGGAAGCAATGTTGGATGAAATTC 2822

RESULT 3
 ID AA292645 standard; cDNA; 2170 BP.

XX AA292645;

XX 05-JUN-2000 (first entry)

XX Murine alpha-1,3-fucosyltransferase cDNA.

XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentialator; diagnosis;
 KW treatment; cancer; murine; mouse; ss.

OS Mus sp.

XX Key Location/Qualifiers

XX CDS 115..1194 /tag="a
 /product="Murine alpha-1,3-fucosyltransferase"

XX WO200006708-A1.

XX 10-FEB-2000.

XX 29-JUL-1999; 99WO-0P04092.

XX 29-JUL-1998; 98JP-0213823.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PI Narimatsu H, Kudo T, Sasaki K;
 XX WPI; 2000-183120/15.
 DR P-PSDB; AAY80995.
 XX
 PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT epitope and an antibody recognizing it useful for diagnosis of brain
 PT and kidney disease and cancer.
 XX
 PS Claim 4: Page 127-134; 172pp; Japanese.
 XX
 CC The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-IV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA92645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA90995), and sequences AA92670-92673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).
 CC
 CC
 XX
 SO Sequence 2170 BP; 715 A; 395 C; 368 G; 692 T; 0 other;
 Query Match 45.1%; Score 1206.8; DB 21; Length 2170;
 Best Local Similarity 78.4%; Pred. No. 3.8e-260;
 Matches 1674; Conservative 0; Mismatches 387; Indels 75; Gaps 16;

QY 187 CCCATGACGCGATTTAGATGATATGATCAAGGATTTGATATGATAGATAGATA 246
 DB 9 CGCTGGGATGCTTTAGAAATGATTAACATCAGCATCTGATAGACCGGTGGAGTAGTTCA 68
 QY 247 ACAACTGTACGTCTCCCATGATATGTTCTATATATGAAAAATATGATCATCA 306
 DB 69 GCACCTACATC - TGCTTCACATCTATGTTCTCTACACGAAAAATATGATCATCA 126
 QY 307 TCACAAGGAATTCCTGCCATTTTAAATGTCGATATCCGGGCTGTTTATGGA 366
 DB 127 TCACAAGGATTCCTGCCATTTTCTATGTCGATCATCTCGGGCTGTTTATGGA 186
 QY 367 TGTCTTCGATTTACATCAACATCAACAGAGTGGATCTCATTCGAATGGATAGCC 426
 DB 187 TGTCTGCTCATTTATATCAACCCACACAGCTGGGCTTCACTCCATGATGATGCA 246
 QY 427 AGCTGTGCTGAAATATGAAAACTTCTTTTCCACAAAAGTGATATTTTAAAGAACT 486
 DB 247 AGTCTGTGCTGAAATATGAAAAATTTCTCTCCACAAAAGTGATATTTTAAAGAACT 306
 QY 487 ACTATTTCTGCTGGGTGGGCTATTTGGGAGACCTTTGACCTTACCTCTGCCAAGA 546
 DB 307 ACCATTTCTGGGTGGGTGGGCTATTTGGGAGACCTTTGACCTTACCTCTGCCAAGA 366
 QY 547 ATGTTCAACATCAAGGATGCTCATCAACGAGCGGTGATGCTCAACAAATCCCAT 606
 DB 367 ATGTTCAATATCCAAAGGTGCTCATCAACAGACCGCTCATGTTTACAAATCCCAT 426

QY 607 GCACTTGATCATCAACGAGACATGATGGATCTGACAAATTTACCTACGAAGCT 666
 DB 427 GCGTCTGATTCACATAGACATGAGATCGGATCTGACTACTACTACCTACGACAGGCC 486
 QY 667 AGGCCACCTTCCGAAGATGATTTGGATGATTTGGAAATACCAACTCACACTCCCA 726
 DB 487 AGGCCACCTTCCGAAGATGATTTGGATGATTTGGAAATACCAACTCACACTCCCA 546
 QY 727 AAGATGGCATTTGAGCACTTTGATTAACCTGACTGATACCGCGTGATTCAGATATC 786
 DB 547 AAGATGGCATTTGAGCACTTTGATTAACCTGACTGATACCGCGTGATTCAGATATC 606
 QY 787 CAACTGCTTATGAGCTTCTTGACGCGTAAGCAAAATCCCTTCTGTTTGAAGTCCAA 846
 DB 607 CAACTGCTTATGAGCTTCTTGACGCGTAAGCAAAATCCCTTCTGTTTGAAGTCCAA 666
 QY 847 AAAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
 DB 667 AAAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
 QY 907 TATTCATATGAGCTATACCAAAAGCATTAATCCATACCTACCGGCAAGATTTGAGAA 966
 DB 727 TATTCATATGAGCTATACCAAAAGCATTAATCCATACCTACCGGCAAGATTTGAGAA 786
 QY 967 TATTCATATGAGCTATACCAAAAGCATTAATCCATACCTACCGGCAAGATTTGAGAA 1026
 DB 787 TATTCATATGAGCTATACCAAAAGCATTAATCCATACCTACCGGCAAGATTTGAGAA 846
 QY 1027 GAAATTCATATCCCAAGGATTTACATCAGGAAAGCATTAATCCATACCTACCGG 1086
 DB 847 GAAATTCATATCCCAAGGATTTACATCAGGAAAGCATTAATCCATACCTACCGG 906
 QY 1087 TCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
 DB 907 TCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
 QY 1147 TCATTCATATGAGCTATACCAAAAGCATTAATCCATACCTACCGGCAAGATTTGAGAA 1206
 DB 967 TCATTCATATGAGCTATACCAAAAGCATTAATCCATACCTACCGGCAAGATTTGAGAA 1026
 QY 1207 GTCGACAAAACATTAAGTATACCTTACCTTATTAACCTGAGAGAGATTTACCTGTA 1266
 DB 1027 GTTTCGACAAAACATTAAGTATACCTTACCTTATTAACCTGAGAGAGATTTACCTGTA 1086
 QY 1267 AATCTTCACGATTTTGGGAAATTCATGATGATGCTGCTGCTGCTGCTGCTGCTG 1326
 DB 1087 AATCTTCACGATTTTGGGAAATTCATGATGATGCTGCTGCTGCTGCTGCTGCTG 1146
 QY 1327 CAAGAATATAGTCTGTTGTAATTTAGAGAAATGTTTGGAAATTTTAAATTTTCAATCA 1386
 DB 1147 CAAGAATATAGTCTGTTGTAATTTAGAGAAATGTTTGGAAATTTTAAATTTTCAATCA 1205
 QY 1387 CTTCGACACTTGAATATTTTATGATGATATATCATCAAGTATTTAGAGATTAAGAGAGA 1446
 DB 1206 TTTCGACACA - GTGAAAAAGATTAATGATGATATTAATCAATTTTCAAGATTAAGAGAGA - 1260
 QY 1447 TGCACATATCTACTTTTGTGTCACAAATTTATTTATCACCTGCTGATAGGTAAGCTGA 1506
 DB 1261 - - - AACACGCTGCAATTTGGGCGACCGTTTATTTTCTGCTGCTGCTGAGACCATGTA 1317
 QY 1507 TATTTTGGTGGAGATTTTAAAGCTCAGATGACATCAATTCATTCGTTGTTTAAAT 1566
 DB 1318 TATTTTGGT - GAATTTTAAAGATATGAGATTTGACATATCTCAATTTGTTTAAAT 1376
 QY 1567 ATCTGTATATATCTAATATATGATGATGATGATGATGATGATGATGATGATGATG 1626
 DB 1377 ATCTGTATATATCTAATATATGATGATGATGATGATGATGATGATGATGATGATG 1435
 QY 1627 AACATGCTTTTACATTTTGTATGTCATTAATGTAAGCTTGGGTTGTTTATG 1686
 DB 1436 AACATGCTTTTACATTTTGTATGTCATTAATGTAAGCTTGGGTTGTTTATG 1489


```

OY 1687 TTCCACACTGATGACGCTTTATCTATTGGGAATGAATGACACTTAAAGT 1746
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1490 TTCTTCATCATCATCATCTTTTATCTATTGGGACATTTAAATCATCTCAAAAT 1549
OY 1747 GA-AAATTTTCACTAGTATTACAAATGCTGTTCCAACTTGCATTAACAGAG 1805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1550 GAGGGAGCTTTGGCAATTTAAAGTATGATTTAGCTTATATGCTTAACAAG 1609
OY 1806 AAGAA--CATGTGGCATTGAATCTAACCTTTGACTCTCAATGATGAAGTGTAT 1863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 AAGAAATTTGATGATCATCTTAATCTTAATCTTGAATCTCAAACTGAACTAGC 1669
OY 1864 AACTGTCTATTTGATCTATTTTAACTGTTATACATTTGTGGAAGGTGAATTA 1923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1670 AACTG-----ATAGAGTTTACTTGTACACATTTGTGAAGGAGACTTA 1715
OY 1924 TTCAATGAGTGAATAGAAAGATATGAACAGACTGTCTATTCAGGAAGCTATTAGC 1983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1716 TTGATGTACCAATAGAAAGATATGAACCAAAATGT-----TATTAGC 1761
OY 1984 TTCTCATTTATTTCTGATTTAGCTGATTTGCACTACTATTTCTCATGCTTAATTA 2043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1762 TTCACTTATTTTATTTATGATTTTACCAATTTGATTTCCAAATTTTAACTTAA 1821
OY 2044 TTATTAGTATTTTAAATATCAATTTGTTGTGAT-----TTCAGC 2087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1822 GAAACAGAAATTTTCAGATTAACATTTATATGTTTAAAGGAGATTTGATTTGTT 1881
OY 2088 ACCGTGGAATATTCCTCAATTAATCTTAAAGAAATC-TAAGACATTTCTTCTGCTAG 2146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1882 TTGTGGGTTAGTCTCTTAATTAATCTTAAGAAATCTTTGATTTGCTTCAAAATCA 1941
OY 2147 ATGACATCATTTGCTATTAATAAACAATAATTTCTCAATTAACAAAGAAAA----- 2200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1942 TGAAGATTAACCACTATTTTAAATAAATAATTTCTCTTATTAACAAATAAATAATGATA 2001
OY 2201 -TGATACCTATTAATATATTTTAAATGATGCTATTATGACATGTTTATATGAT 2259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2002 TCTTCAAAATACAAATTAATGATTAATATATACATTTATAGAACCTTGTACT-TCAAT 2060
OY 2260 CAATTAAGATTTTCTTGAAGCCCTAATTTAA 2295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2061 CATTTGAAATATTTTTCACAAAGCATCATTTAA 2096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4
AA292672 standard; cDNA, 2036 BP.

AA292672;

05-JUN-2000 (first entry)

Murine alpha-1,3-fucosyltransferase gene exon 1 DNA, SEQ ID NO:30.

Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
brain; kidney; recombinant expression; transgenic animal; knockout
animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
treatment; cancer; murine; mouse; ss.

Mus sp.

WO200006708-A1.

10-FEB-2000.

29-JUL-1999; 99WO-JP04092.

29-JUL-1998; 98JP-0213823.

(KIOM) KYOMA HAKKO KOYO KK.

Narimatsu H, Kudo T, Sasaki K;

XX WP: 2000-183120/16.
DR Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
XX epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer.
PT
PT
XX
XX
PS
PS

Example 8: Page 162-164; 172pp; Japanese.

The invention relates to a novel alpha-1,3-fucosyltransferase which
transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
(generating the Lewis x or y epitope). It does not transfer a fucose
moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
and therefore does not generate the sialyl-Lewis x epitope. The
invention also relates to DNA sequences encoding alpha-1,3-
fucosyltransferase and expression vectors and host cells comprising
these DNA sequences. The invention additionally encompasses the
preparation of alpha-1,3-fucosyltransferase via the culture of
transformed cells or by expression of the protein in a transgenic animal;
screening potential inhibitors or potentiators of
alpha-1,3-fucosyltransferase activity or expression; the preparation of
compounds having fucose-containing sugar chains by use of the protein;
and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
Alpha-1,3-fucosyltransferase has a similar substrate range to the known
FUC-TIV and is expressed mainly in brain and kidney tissues.
Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
potentiators and inhibitors may be used for the treatment and diagnosis
of diseases of the brain and kidney, and of cancers. They may be used for
the identification of substances which affect the activity or expression
of alpha-1,3-fucosyltransferase; such substances may be used
therapeutically. The knockout animals can be used to study the mechanisms
of action and expression of alpha-1,3-fucosyltransferase. Sequence
(AA292645) represents cDNA encoding murine alpha-1,3-fucosyltransferase
(AA292670-292673) are murine
alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
amplification of cDNA ends).

Sequence 2036 BP; 666 A; 365 C; 348 G; 657 T; 0 other;

Query Match 43.3%; Score 1159.2; DR 21; Length 2036;
Best Local Similarity 78.6%; Pred. No. 1.6e-249;
Matches 1600; Conservative 0; Mismatches 363; Indels 73; Gaps 15;

```

OY 287 GAAAAATTTATGACATCAACATCCAAAGAAATTTCTGGCCATTTTAAATGCTGCATTA 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 GAAAAATTTATGACATCAACATCCAAAGCAATTTCTGGCCATTTTAAATGCTGCATCA 60
OY 347 TCTGGGCTGTTTCATGATGCTCTTCTCATTTACATCAAACTTACCAACAGCTGATCT 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 TCTGGGCTGTTTCATGATGATGCTCTCTCATTTATTAACAAAGCCCAACAGCTGATCT 120
OY 407 TCAATCCAAATGATGACAGCAGCTCTGCTGTAATAATGAAAACTTTTCCACCAAAAA 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 TCAATCCAAATGATGATGACAGCTCTGCTGTAATAATGAAAACTTTTCTCTCAAAAA 180
OY 467 CTGATTTATTTTAAATGAATGATTTCTGTTGGGCTGCTGATTTGGGCAAGCTTTG 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 CTGATTTATTTTAAAGCAATCATCTCTGTTGGGATGATGCGCATTTGGGCAACCTTTG 240
OY 527 ACCCTTACATCTGCCAAGCAATGTTCAACATCCAAAGATGCTGCATCAAAAGACGCT 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 ACCCTTACATCTGCCAAGCAATGTTCAATATCCAAAGGATGCTGCATCAACAGACGCT 300
OY 587 CACTGTCAACAAATCCCAATGCAATTTGATCATCAACGACATCATGTTGGATCTCA 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 CATTTGATCAACAAATCCCAATGCAATTTGATCATCAACGACATCATGTTGGATCTCA 360
OY 647 CAATTTACCTCAAGCAAGTAGCCACACCTTCCAGAAATGATTTGATGAATTTGGAAT 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 CTAACTTACCTCAAGCAAGTAGCCACACCTTCCAGAAATGATTTGATGAATTTGAGAT 420
OY 707 CACCACTCAAGTAGCCCAAAAGATGAGCATTTGATGAGCATTTGATGAGCATTTGACT 766

```



```

Db 421 CACCCCTCACACCCCAAGAGAGGACATGACATGCTGTCACCTGCTACTCTACTT 480
Qy ACCCGGTGATTCAGATATCCAAAGTGCCTTATGGCTTCTTGACGTAAAGCAAAATCCCT 826
Db 481 ATGCCGTGATTCAGATATCCAAAGTGCCTTATGGCTTCTTGACGTAAAGCAAAATCCCT 540
Qy TCGTGTTCAGATTCAGCAAGCAAGAAATGCTGTGCTGGTGTGAGTAACTGGAACC 886
Db 541 TTGTGTTCAGATTCAGCAAGCAAGAAATGCTGTGCTGGTGTGAGTAACTGGAACC 600
Qy CTGAGATGCGCAGATCAATATTAATAGCTAAAGCAAAAGCAATTAATTAATCAATCT 946
Db 601 CTGAGATGCGCAGATCAATATTAATAGCTAAAGCAAGCAAGCAATTAATTAATCAATCT 660
Qy ACGGCAAGCATTTGGAGATATGTCATATTAATTAATTAATTAATTAATTAATTAATCT 1006
Db 661 ATGCCCAAGCATTCGGAATATGCTGACGATTAATTAATTAATTAATTAATTAATTAATCT 720
Qy GTAAATTTATCTTCTTCTGAAATTCACCAAGATTAATTAATTAATTAATTAATTAATCT 1066
Db 721 GTAAATTTATCTTCTTCTGAAATTCACCAAGATTAATTAATTAATTAATTAATTAATCT 780
Qy ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
Db 781 ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy AGAATTTATTCAGAGATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1186
Db 841 AGAATTTATTCAGAGATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 900
Qy TACCAAAATATCTGAAAGAGTTCGACAAAACAATTAATTAATTAATTAATTAATTAATTA 1246
Db 901 TACCAAAATATCTGAAAGAGTTCGACAAAACAATTAATTAATTAATTAATTAATTAATTA 960
Qy GGAGGAAGCATTCACCTGATTAATTCCTCAGATTTGGGAAATGACATGCTGCTGCT 1306
Db 961 GGAGGAAGCATTCACCTGATTAATTCCTCAGATTTGGGAAATGACATGCTGCTGCTGCT 1020
Qy GCGATCTGTAAGAAAGCATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1366
Db 1021 GCGATCTGTAAGAAAGCATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Qy GGAATTTAAATTTTTCATCCTTGCACACTTGATTAATTAATTAATTAATTAATTAATTAAT 1426
Db 1081 GGAATTTAAATTTTTCATCCTTGCACACTTGATTAATTAATTAATTAATTAATTAATTAAT 1136
Qy GTATTCAGATTAAGAGATGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1486
Db 1137 GTATTCAGATTAAGAGATGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1191
Qy CCTCTGAGGTAAGGTAATTTGGTGAATTTTAAAGCTGACATGAGCAATC 1546
Db 1192 CCTCTGAGGTAAGGTAATTTGGTGAATTTTAAAGCTGACATGAGCAATC 1250
Qy ATTCAATTCGCTTTAAATTAATTCCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1606
Db 1251 ATTCAATTCGCTTTAAATTAATTCCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1309
Qy ATTCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1666
Db 1310 ATTCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1366
Qy AGCTGTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1726
Db 1367 AGCTGTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1423
Qy AGATGACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1785
Db 1424 AGATGACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1483
Qy TTTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1843

```

```

Db 1484 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1543
Qy 1844 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1903
Db 1544 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1589
Qy 1904 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1963
Db 1590 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1647
Qy 1964 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2023
Db 1648 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1695
Qy 2024 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2081
Db 1696 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1755
Qy 2082 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2126
Db 1756 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1815
Qy 2127 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2186
Db 1816 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1875
Qy 2187 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2239
Db 1876 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1935
Qy 2240 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2295
Db 1936 CTTCATTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1990

RESULT 5
AAT59505
ID AAT59505 standard; cDNA; 1814 BP.
XX
XX AAT59505;
AC
AC AAT59505;
DT 06-MAY-1997 (first entry)
XX
XX Murine myeloid-lineage alpha-(1,3)-fucosyltransferase cDNA.
DE
XX Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM;
KW septic shock; septicemia; therapy; ss.
XX
XX Mus sp.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 325..1353
FT CDS /*tag a
XX
XX WO9640881-A1.
XX
XX 19-DEC-1996.
XX
XX 08-MAY-1996; 96MO-US06427.
XX
XX 07-JUN-1995; 95OS-0483151.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX
XX Holgersson J, Seed B.
XX
XX MPI; 1997-108639/10.
XX
XX P-PSDB; AAW11820.
XX
XX New murine alpha-(1,3)-fucosyltransferase - for fucosylating an
XX antibody to protect mammals against e.g. septic shock or septicemia
XX
XX Claim 6; Fig 3; 58pp; English.

```


CC production of oligosaccharides, to generate antibodies to
CC glycosyltransferases useful as diagnostic reagents, to screen for
CC fucosyltransferase inhibitors and inactivators, especially those
CC that act as antiinflammatory agents, for genotyping individuals at
CC the fucosyltransferase locus and for in vitro synthesis of sialyl
CC Lewis x tetrasaccharide.

XX
SQ Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other;

Query Match 4.7%; Score 127; DB 18; Length 3594;
Best Local Similarity 52.3%; Pred. No. 1,9e-18;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

OY 563 GATGACATCTACACAGGACCGTTCACTGTACAAACAATCCAGTGTGATCATC 622
DB 2277 GCTGCGCTGTAGTGTAAACCGAGCGCTGTACAGTGTGTGCTGTTCACAC 2336
OY 623 ACCGACATCACTGGGATCTGACAAATTACTCAGCAGCTAGCCCTTCACAG 682
DB 2337 ACCGTAGCTGCAAAACCCGCAATCTCTCTACCCCTGACACAGGCCACAGGAGC 2396
OY 683 AATGATTTGGATGAAATTTGGATCAACACACACACCTCCCAAAAGAGGCGATTGAGC 742
DB 2397 CTGGGCTGTGGCTCCATGGAATCCGCCAGTAATACCAATGCTTCATCGCTCCGG 2456
OY 743 ACTTGTTTAACTGACTGTACTTACCGCGGTGATTCAGATATCAAGTCCCTTATGCT 802
DB 2457 GCATCTTCAACGTGGTGTGAGCTATCGGGGTGATTCAGATATCTTGTACCCCTAGGTC 2516
OY 803 TCTGACGCTAGCAACAATCCCTTGTGTGTGTTTCAAGTCCAGCAAAAGAGAAATGTG 862
DB 2517 GCTTGGAGCCCTCTCTGTGGCC--CACAATCCCACTACCGGCAAAAGAGATGCTG 2573
OY 863 GCTGGGTGTGAGTACCTGGAACCTGAGATCCAGAGTCAAGTAATACATGAGCTAA 922
DB 2574 CCGGAGTGTACCAATTTTCCAGAGCGGACGACAGCTGTACCGGACGCTGG 2633
OY 923 GCAAAAGCATTTGAATCATACCTACGAGCGCAGCATTTGGAGAAATGTCAATGATAAA 982
DB 2634 CCCCTCATCTGCAAGTGTGTGTGCGCGCCAGCGAGCGGCCCATATGGCTAAT 2693
OY 983 ATTGATTTCTACATATCTGCTTGAATTTATCTTTCTTTGAAATTCATCCACA 1042
DB 2694 GTCTGTGCGCCACTTTGGCCCGGCTACCGCTTACCTGGCTTTGAAACTCAGACATC 2753
OY 1043 AGGATTAACATCAGGAAAGCTAT--ACAATGCTTTTCTGSGCTGTACTCGTTG 1099
DB 2754 GGGACTACATCACTGAAAGTCTGTGCGCAATCCCTGCGGCTGTGTGATCCCGTGG 2813
OY 1100 TTCTGGACCATCTAGGAAACTATGAGAAATTAATTCAGAGATTCATTCATG 1159
DB 2814 CGCTGGAGCTCTCTGGGCGACCTACGAGGCTTTTGTGCACAGAGTCTTTGTACAG 2873
OY 1160 TGGAGATTAACTCTCCAGTGAAGCTAGC 1190
DB 2874 TGGACGACTTCAGCTCTGCCGCTGAAGTGGC 2904

RESULT 7

ID AAQ77732 standard; cDNA to mRNA; 1701 BP.

XX AAQ77732;

XX 22-JUN-1995 (first entry)

DE Human alpha-1,3-fucosyltransferase cDNA.

XX Alpha-1,3-fucosyltransferase; sialylated Lewis blood group antigen;

KW Lewis x; Lewis a; glycosylation; glycoprotein; glycolipid; ds.

XX Homo sapiens (THP-1 cells).

EH Key Location/Qualifiers
FT CDS 123..1151
FT /tag= a
FT /product= alpha-1,3-fucosyltransferase

PN W09423021-A.

XX 13-OCT-1994.

XX 28-MAR-1994; 94WO-JP00496.

XX 29-MAR-1993; 93JP-0069016.

XX (KYOW) KYOMA HAKKO KOYO KK.

XX Hanai N, Kurata K, Nishi T, Sasaki K;

XX WPI; 1994-333183/41.

XX P-PSDB; AAR63215.

PT Alpha-1,3-Fucosyl-transferase DNA and protein - useful for prodn.

PS of active sugar chains, e.g. sialylated Lewis x

XX Claim 3; Page 99-102; 155pp; Japanese.

CC The sequence coding for alpha-1,3-fucosyltransferase was isolated
CC from a human THP-1 cell cDNA library. The enzyme catalyses the
CC addition of fucose via a 1,3-linkage to N-acetylglucosamine
CC contained in glycoproteins and glycolipids, e.g. for sialylation of
CC Lewis blood group antigens.

XX
SQ Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;

Query Match 4.3%; Score 114.4; DB 15; Length 1701;
Best Local Similarity 50.8%; Pred. No. 9.4e-16;
Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

OY 550 TTCAACATCCAGAGATCCATCTCACACGACCGTTCACTGTACAAACAATCCATGCA 609
DB 333 TAGGCGATGCCCGCGTGCACCTGTAGTCCACAGCAAGCTGCTGGCCAGCGGACGCC 392
OY 610 GTTCTGATCCATCAGCGAGCATCACTGGGATCTGACAAATTTACTCAGCAAGCTAG 669
DB 393 GTGTGTCTTCCACCAACGCGAGCTGACAGCCCGCGGTCCCACTGCGCGGCCAGCG 452
OY 670 CCACCTTCCAGAAATGGATTTGGATGATTTGGATTCACCACTACACTCCCAAAAG 729
DB 453 CCGGAGGAGGACGCGTGGGTGGGCTCCATGAGTCTCTAGCCACACCGGCGCTC 512
OY 730 AGTGGCATTTGAGCACTGTTTAACTGACTGTACTTACCGCGCGGTGATTCAGATTC 789
DB 513 AGCACCTCCGAGGATCTTCAACTGTGTGTAGTACCGGCGGCGACATGATCTTT 572
OY 790 GTGCTTATGCTTCTTGAACGCTAGCAGCAAAATCCCTGCTGTTTGAATGCCAAGCAA 849
DB 573 GTGCCCTATAGCGCGCTGTGAGGCCCACTGGGGGCCCT--CGCACCGCTGCCAGCAA 629
OY 850 GAGAAATGTGTGTGGGTGTGTGAGTAAGTGAACCCGAGATCCGAGATCCAGATGAT 909
DB 630 AGCAGGATGCGCGCTGGGTGTGTGAGTAAGTGAACCCGAGATCCGAGATCCAGATG 689
OY 910 TACAAATGAGCTTAAGCAAAAGCATTTGAATTCATCTAGCGGCAAGCATTTGGAGAA 969
DB 690 TACCGGAGCTGTGGCGCTCATCTGCGGATGTGTGCTTTGGCGCGTGCATGAGACGCC 749
OY 970 GTCAATGATTAATAATTTGATCTTCACTATCTGCTTGAATTTTATCTTCTTTGAA 1029
DB 750 CTGTGGCGCAGCTGTGTGGGTGTGTGAGTACCGGCGGCGGCGGCTTCACTGTCTTT 809
OY 1030 AATTCAATCCAGAAAGATTAACATCAGGAAAGCTAT--ACAATGCTTTCTGGGTGG 1086
DB 810 AACTCTACGACCGGCGATTAATTAAGGAAATTTCTGGCGCAAGCAAGCTGTGGCTGG 869

KM allergic disease; bronchoconstriction; inhibitor; antinflammatory;
KM antiallergic; antiallthmatic; cyostatic; analgesic; impaired airway;
KM lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KM respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KM pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KM cancer; leukemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI: 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure: Page 957-958; 1343pp; English.
XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiallthmatic, cyostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32313 to AAA3992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.
XX
SQ Sequence 1701 BP: 292 A; 524 C; 586 G; 299 T; 0 other:
XX
Query Match 4.3%; Score 114.4; DB 21; Length 1701;
Best Local Similarity 50.8%; Pred. No. 9.4e-16;
Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

QY 730 AGTGGATTGAGCATTGTTTAACCTGACTGACTTACCGCCGTGATTCAGATATCCAA 789
DB 513 AGCCACTCGAGGACATCTCAACTGAGTCTAGTACCGGGGACATCGACATCTT 572
QY 790 GTCCCTTATGCTCTTGTGACGGTTAGCACAATCCCTGCTGTTTAAAGTCCAGCA 849
DB 573 GTCCCTTATGCTCTTGTGACGGTTAGCACAATCCCTGCTGTTTAAAGTCCAGCA 849
QY 850 GAGAAATTTGCTGTGGTGTGAGTACCTGGAACCTGAGATCCAGATCAATAT 909
DB 630 AGCAGGATGCGCCCTGTGGGTGTGAGTACCTGGAACCTGAGATCCAGATCAATAT 909
QY 910 TACATGAGCTTAAGCAAAAGCATTTGAATTCATCTAGCGCAAGCATTTGGAATAT 969
DB 690 TACCGGACAGTGGCGGCTCATCTGCGGGGTGAGTCTTTGGCCGTGCAATGACGCCA 749
QY 970 GTCAATGATTAATAATTTGATTCCTACCAATATCTGCTTTGTAATTTATCTTCTTGA 1029
DB 750 CTGTGCGCAGCTGCTGTGGTGTGAGTACCTGGAACCTGAGATCCATCTTCTTGA 809
QY 1030 AATTCAATCCAGAAAGATTACATCAGGAAAGCTAT--ACAATGCTTTCTGCTGCG 1086
DB 810 AACTCTCAGCAGCGGACTTACATTTACGAGAAATTTCTGCGCAGCAGCTGTGCTGCG 869
QY 1087 TCTGTACCTGTTGTTCTGAGGACCATCTAGGAAATATGAGAAATTTATTTCCAGCAGAT 1146
DB 870 ACTGTGCAAGTGTGCTGTGGGGCCCCACGGGCCACCTATGAGGCTTCTGCGGCTGAC 929
QY 1147 TCATTCATTCATGTCGAGATTTAATCTCTCCAGTGAAGCTAGC 1190
DB 930 GCCTGTGTCATGTGATGATCTTTGGCTCAGCCGAGAGCTGCG 973
RESULT 10
ID AAF21137 standard; DNA; 6944 BP.
XX
AC AAF21137;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #2704.
XX
KM Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KM human; airway disorder; bronchoconstriction; lung inflammation;
KM surfactant depletion; respiratory; bronchodilator; antinflammatory;
KM immunosuppressive; antiallthmatic; analgesic; hypotensive; cyostatic;
KM respiratory obstruction; pulmonary obstruction; impeded respiration;
KM surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KM respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KM pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KM chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KM cancer; ss.
XX
OS Homo sapiens.
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08020.
XX
PR 06-APR-1999; 99US-0127958.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PA (NYCE/) NYCE J W.
XX
PI Nyce JW;
XX
DR WPI: 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not

DE Lewis blood group fucosyltransferase DNA.

KW Glycosyltransferase: fucosyltransferase: GDP-Fuc: in vitro; cell;

KW surface; oligosaccharide; ss.

PN MO9402616-A.

XX 03-FEB-1994.

PD 20-JUL-1993; 93WO-US06703.

XX 20-JUL-1993; 93WO-US06703.

XX 20-JUL-1992; 92US-0914281.

XX (UNMI) UNIV MICHIGAN.

PI Lowe JB;

DR WPI; 1994-048874/06.

XX DNA fragment encoding a glycosyltransferase - can be used for in

PT vitro reactions to modify cell surface oligosaccharide(s) e.g.

PT blood gp. determinants, to protect against transplant rejection

XX

PS Disclosure; Fig 6; 24pp; English.

XX The sequence is that encoding a Lewis blood group fucosyl

CC transferase (Fuc-TIII). The enzyme produced by the DNA may be

CC non glycosylated. This prevents premature loss of enzyme activity.

CC It can also be used in in vitro reactions to modify cell surface

CC oligosaccharide mols. e.g. blood group determinants.

CC See also A056905-11.

XX

XX Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;

XX

Query Match 4.2%; Score 112.2; DB 15; Length 1086;

Best Local Similarity 51.6%; Pred. No. 2.4e-15;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 565 TGCCATCTCACAGCAGCGCTTCACTGACAAATCCATGAGTTCGATCCATC 624

DB 271 TGCACATCACTGCGCGAGCGAGGTGACCAAGGAGAGAGGATCGTGACAC 330

QY 625 CGAATCATCACTTGGGA---TCTGACAAATTACTCAGCAGCTAGGCCACCTTCAG 681

DB 331 TGGATATCACTGACCACTCAAGCTCAGCCTCCACCTCCGAGGCCAGGGGAG 390

QY 682 AAATGATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 741

DB 391 CGCTGATCTGCTTCACTTGGAGCCACCCCTTAAGTCCAGCAGCTGGAAGCCCTG 450

QY 742 CACTTGTAACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 801

DB 451 AGATATCTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 510

QY 802 TTTCTGACGGAAGCACAATCCCTTCTG-----GTTTGAAGTCCCAAGCAAGAG 852

DB 511 TGGCTGGAGCCGTGCTGCGGCGAGCTGCCACCCAGCTCAACCTCTCGCCAAAGCC 570

QY 853 AAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912

DB 571 GAGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 630

QY 913 AATGAGCTAACCAAAAGCATTTGAATCCATCTACAGGGGCAAGCATTTTGGAATATGTC 972

DB 631 CAGAGCCCTGAGGCTCATCTCAAGGTGAGCTGAGAGCTGCTCCACAAGCCCTGCGCC 690

QY 973 AATGATTAATAATTTGATTTCTTCACTATCTGCTTGAATTTTATCTTTCTTTGAAAT 1032

DB 691 AAGGGGACCA---TGATGAGACAGCTGCTCCGAGTCAAGTTCTACTGCTTCCAGAGAC 747

QY 1033 TCAATCCACAAAGATTAATCAACGGAAGAAAGTAT---ACAATGCTTTCTGGCTGGCTCT 1089

DB 748 TCTCTGACCCCGACTCATCACTCCAGAGAGCTGTGAGAGAACCCCTTGGAGGCTTGGGCC 807

QY 1090 GTACCTGTTGTTCTGGGACCATCTAGGAAACATATGAAATTATATTCAGCAGATTCA 1149

DB 808 GTGCCCGGTGCTGGGCGCCAGAGCAACTACGAGAGATTCTTCCACCCGAGCC 867

QY 1150 TTCAATTCATGTGGAAGATTAATACCTCCAGTGCATGACGAAGTATCTGAGAGCTC 1209

DB 868 TTCATCCAGCTGAGAGATTCCAGAGCCCCCAAGACCTGGCCCGGTACTCTCAGAGACTG 927

QY 1210 GACAAACAAATTAATTAATCTAGTACTTAACTGAGAGAG 1254

DB 928 GACAAGACCAAGCCGCTACTGACTGACTTTCGCTGCGGGAG 972

RESULT 15

AAT80111

ID AAT80111 standard; cDNA; 2042 BP.

XX

AC AAT80111;

XX

DT 01-DEC-1997 (first entry)

XX

DE Fucosyltransferase III coding sequence.

XX

KW P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;

KW calcium-dependent carbohydrate binding protein; platelet; endothelium;

KW thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;

KW inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;

KW disseminated intravascular coagulation; rheumatoid arthritis; antibody;

KW adult respiratory distress syndrome; tumour metastasis;

KW fucosyltransferase III; acetylglucosaminyltransferase;

KW autoimmune disorder; inflammatory disorder; ss.

XX

OS Homo sapiens.

XX

PN W09706176-A2.

PD 20-FEB-1997.

XX

PF 02-AUG-1996; 96WO-US12820.

XX

PR 17-MAY-1996; 96US-0649802.

PR 03-AUG-1995; 95US-0510920.

PR 13-MAY-1996; 96US-0017794.

XX

PA (OKLA) UNIV OKLAHOMA STATE.

XX

PI Cummings RD, McEver RP, Moore KL;

XX

DR WPI; 1997-154206/14.

XX

PT New O-glycan(s) derived from P-selectin glyco:protein ligand 1

PT (PSGL-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.

PT for treating inflammatory or autoimmune disorders or tumours

XX

PS Disclosure; Page 80-81; 96pp; English.

XX

CC This sequence represents the coding sequence for the human

CC fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-

CC acetylglucosaminyltransferase encoded by AAT80112 are needed for the

CC expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in

CC mammalian cells. P-selectin is a calcium-dependent carbohydrate binding

CC protein expressed on the surfaces of activated platelets and endothelium

CC in response to thrombin and other agonists. PSGL-1 (see A06174) is a

CC high affinity P-selectin ligand produced by leukocytes. Binding of

CC P-selectin to PSGL-1 is calcium ion dependent and is abolished by

CC treatment of the ligand with sialidase. PSGL-1 is a homodimer, with a

CC highly extended extracellular domain, which is a feature of mucin-like

CC proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of

CC the invention (see A06175-W26188). PSGL-1 is also targeted by the

CC O-glycans, and O-glycanated products of the invention. The products are

CC used for inhibiting the binding of P-selectin and other selectins to

CC PSGL-1. They can be used for the treatment of inflammation, injury

Tue Oct 8 10:18:31 2002

us-09-744-748-5.rng

Page 16

CC resulting from ischaemia and reperfusion, bacterial sepsis and
CC disseminated intravascular coagulation, adult respiratory distress
CC syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and
CC other autoimmune or inflammatory disorders. The products can also be used
CC in the production of antibodies and in diagnostic applications.
XX

SQ Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

Query Match	4.28;	Score 112.2;	DB 18;	Length 2042;
Best Local Similarity	51.6%;	Pred NO 3	1e-15;	

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY	565	TGCGATTCACAAAGGACCGCTTCACGTGACAAACAATATCCATTCAGTTGATCAATCAC	624
Db	343	TGCACATCACTGCGCGACCCGAAGGTATACCACAGCAGACACGGTCAATCGTCGACACAC	402
QY	625	CGACATCATAGTTGGGA---TGTGACAAATATTCACACAGACACTTGGGCAACCTTCGAC	681
Db	403	TGGGATATCATGTCCAACCCCTAAAGTACAGCGCTCCACACTTCCCCGAGGGCGGAGGGCGAC	462
QY	682	AAATGGATTTGGATGATTTGGATATACCAACATCAACATCCCAAAAGATGGCAATTGAG	741
Db	463	CGCTGATCTGGTTCAACTTGGAGCCACCCCTTAATCGCAGACACTGGAAACCTCTGACAC	522
QY	742	CACATTGTTAACGATCTGCATCTACCTACCGCGCGTGATTCAGATATCCAAAGGCTTAATGGC	801
Db	523	AGATACTTCATCTACCAATATGTCTACACGAGGAGACTCCGACATCTTACCGGCTTAAGGC	582
QY	802	TTCTTGACGGTAAAGCAAAATCCCTTGGT-----GTTTGAAGTGGCAGCAAAAGAG	852
Db	583	TGGCTGAGCGCTGTGTCCGGCCAGCGCTCCCAACCCAGCTCAACCTCTCGGCCCAAGAC	642
QY	853	AAATGGTGTGCGGGTGTGGATGATCTAGAACCCCTGACATATCCGAGATCAAAATATAC	912
Db	643	GAGCTGTGGCGTGGGGGTGTCCAACTGGAAACCGGACTCAGCCAGCGGTGGCGCTACAC	702
QY	913	AATGAGCTAAGCAAAAGCAATTGAAATTCATACCTACAGGGCAACGATTTGGACATATATGC	972
Db	703	CAGAGCCTCGAGGCTCATCTCAAGGTGAGACGTATACGACGCTCCACAAAGCCCTGCCC	762
QY	973	AATGATTAATTTGATATCTCTACCATATCGCTGTGAAATTTATTTCTTCTTTGAAAT	1033
Db	763	AAGGGGACCA---TGATGAGAGCGCTGTCCCGGTACAAAGTTCTTACCTGGCTTTCGGAAC	819
QY	1033	TCATCTCACAGAGATTTCATACGAGAAAGCTAT---ACAAATGCTTTTCTGCGTGGCTCT	1088
Db	820	TCTCTGACACCCCACTATCATACCGAGAAACCTGTGGAGAACGCCCTGGAGAGCTGGGGC	879
QY	1090	GTACCTGTGTTCTTGGGACCATCTAAGGAAACTATGAGAAATATATATATTCACAGAGATTCA	1145
Db	880	GTGCGCCCTGTGTGGGGCCCAAGACAGAAAGACTTACGAAAGGTTCTCTGCAACCCGACGCC	939
QY	1150	TTCAATTATGTGAGAGATTATTAATCTCCCACTGAGACTACAAAGTATCTGAGAAAGTC	1205
Db	940	TTCTATTCACCTGGACGAGCTTCCAGAGCCCCCAAGACCTTGACCGGTAACTGTGAGAGCTG	999
QY	1210	GACAAAAACAATAAGTATATACCTTAGTACTTAACTGAGAGAG	1254
Db	1000	GACAAAGGACCAAGCCGCTACCTGAGAGTACTTTCGTTGGCGGGAG	1044

```
Search completed: October 5, 2002, 23:15:38
Job time : 369.303 secs
```

Job time : 369.303 secs

[illegible]

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho., Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
url:http://genome.gsc.riken.go.jp/
Carninci P, Shibata Y, Hayashi M, Sugahara Y, Shibata K, Itoh M, Konno H, Okazaki Y, Muramatsu M, and Hayashizaki Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10): 1617-1630 (2000)
wagi K., Fujiwake S., Inoue K., Toyawa Y., Itawa M., Ohara E., Watabiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai U., Okazaki Y., Muramatsu M., Inoue Y., Kita A. and Hayashizaki Y.
RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno H., Fukunishi Y., Shibata K., Itoh M., Carninci P., Sugahara Y. and Hayashizaki Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo S., Shinagawa A., Salto T., Kiyosawa H., Yamakawa I., Aizawa K., Fukuda S., Hara A., Itoh M., Kawai J., Shibata K. and Hayashizaki Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues

FEATURES

SOURCE

1. 697
location/DnaIifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/cd_xref="taxon:10990"
/clone_cj30004C19
/clone_lib=RIKEN full-length enriched, ES cells"
/cell_type=ES cells
/lab_host="SDSR"
/note="Site:1: XhoI; Site:2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',
GAGGACAGACGATCCACAGACGCTTTTGTGGTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGGACAGACTTCGCGATTAAATTAATTAATCCCCCCCCC 3']".

BASE COUNT 219 a 138 c 141 g 199 t

ORIGIN

Query Match 21.5%; Score 575.6; DB 9; Length 697;
Best Local Similarity 91.2%; Pred. No. le 97;
Matches 622; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

698 ATTGGATCAACAACACTACACTCCCAAAGAAGTGCGATTGAGCATCTGTTAACCCTGA 757
DB 4 AATTAAGATCACCACTACACACCCCAAAAAGAGGCGA-TGACNACTGTGTCAACCTGGA 62
758 CTCGCACCTTACCGCGGTTCAGATATCAATCAAGTCCCTTATGAGGCTCTTGACGGTAACGA 817

Email: dmelton@hlohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 436.

FEATURES

```

source
1. . 579
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/notes="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Site 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies) . cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
BASE COUNT      203 a      103 c      93 g      180 t
ORIGIN

```

BASE COUNT	203 a	103 c	93 g	180 t	.
ORIGIN					

Query Match	21.48;	Score 573.2;	DB 10;	Length 579;
-------------	--------	--------------	--------	-------------

```
Matches 575; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Db	603	GCCTGCATGCGACACGTATAAAGGATCAAGATATATAGTCTGTGGTAAATTATAGACA	662
Qy	1358	AATGCTTTTGGCATTAATAATTT 1379	
Db	663	AATGCTTTTGGAAATTAAGTGT 684	
RESULT 3			
BM021081/c			
LOCUS	BM021081	579 bp	mRNA
DEFINITION	le75f03.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens		
	CDNA 3' similar to TR:Q9Y231 Q9Y231 ALPH-3-FUCOSYLTRANSFERASE. ;		
ACCESSION	BM021081		
VERSION	BM021081.1	GI:16535437	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 579)		
AUTHORS	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradowh,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistah,A., Schmitt,L.A., Theising,B., Ritzer,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y., and Bowers,Y.		
TITLE	Endocrine Pancreas Consortium		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue		
	Harvard University, Howard Hughes Medical Institute		
	Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,		
	MA 02138		
	Tel: 617-495-1812		
	Fax: 617-495-8557		


```
VERSION      B41073.1  GI:2545325
KEYWORDS     GSS.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 501)
             Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
             Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
             Construction of a Characterized Clone Resource for Genomic
             Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
             Tagged Connectors
JOURNAL      Unpublished (1997)
COMMENT      Contact: Mahairas GG, Zackrone KD, Hood L
             University of Washington
             Seattle, WA 98195, USA
             Tel: (206) 616-8744
             Fax: (206) 685-7301
             Email: kzackrone@u.washington.edu
             Sequence Tagged Connector
             Plate: CT 774. row: N column: 2
             Class: BAC ends
             High quality sequence stop: 501.
FEATURES     Location/Qualifiers
             1..501
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="CT 774 Col-2 Row-N"
             /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
             E-Coli DH10B"
BASE COUNT   179 a 95 c 83 g 143 t 1 others
ORIGIN
Query Match      16.7%; Score 448; DB 12; Length 501;
Best Local Similarity 99.8%; Pred. No. 7.7e-74;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1224 GGTATACCTTACTTACTTAACTGAGAGAAAGTTCACCTGTAATCTCCACGATTTG 1283
Db 501 GTATTACCTTACTTACTTAACTGAGAGAAAGTTCACCTGTAATCTCCACGATTTG 442
OY 1284 GGAATCAGACAGCATGTTGGCTGCGATCGATGAGAAAGCAAGATATTAAGCTGT 1343
Db 441 GGAATCAGACAGCATGTTGGCTGCGATCGATGAGAAAGCAAGATATTAAGCTGT 382
OY 1344 TGGTAATTTAGAGAAATGTTTGGAAATTAATTTTTCATCATCTGCACACTTGATTA 1403
Db 381 TGGTAATTTAGAGAAATGTTTGGAAATTAATTTTTCATCATCTGCACACTTGATTA 322
OY 1404 TATTTTGATGAGATATCATCCAGATTTGAGATTAAGAGAGATGCAACATCTACTTTT 1463
Db 321 TATTTTGATGAGATATCATCCAGATTTGAGATTAAGAGAGATGCAACATCTACTTTT 262
OY 1464 GTGTACAATTTTATTTTATCACCCCTCTAGAGGTAACGTGTATTTTGGTGGAGATT 1523
Db 261 GTGTACAATTTTATTTTATCACCCCTCTAGAGGTAACGTGTATTTTGGTGGAGATT 202
OY 1524 TTTAAAGCTCAGATGAGCAATCATCTGGTTTAAATATATCTGTAATATCTTAA 1583
Db 201 TTTAAAGCTCAGATGAGCAATCATCTGGTTTAAATATATCTGTAATATCTTAA 142
OY 1584 TTTATGACACTGAGAGATATTTATCTCATATATATCATTTGTAACATGCTTTTACCA 1643
Db 141 TTTATGACACTGAGAGATATTTATCTCATATATATCATTTGTAACATGCTTTTACCA 82
OY 1644 TTTTGTAGTGTGTCATATATGTAAGCTTG 1672
Db 81 TTTTGTAGTGTGTCATATATGTAAGCTTG 53
RESULT 8
```

```
AV338075
LOCUS       649 bp. mRNA linear EST 24-OCT-2001
DEFINITION AV338075 RIKEN full-length enriched, adult male olfactory bulb Mus
            musculus cDNA clone 6330407824.3, similar to A0015426 Mus musculus
            mRNA for alpha1,3-lucosyltransferase IX, mRNA sequence.
ACCESSION   AV338075
VERSION     AV338075.2  GI:16396626
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM   Mus musculus
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
             1 (bases 1 to 649)
             Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
             Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda
             ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
             Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
             ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
             Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
             Muramatsu,M. and Hayashizaki,Y.
             Riken Mouse ESTs (Arakawa,T., et al. 2001)
             Unpublished (2001)
             On Nov 11, 1999 this sequence version replaced gi:6378127.
TITLE       Yoshinide Hayashizaki
JOURNAL     Laboratory for Genome Exploration Research Group, RIKEN Genomic
COMMENT     Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gs.riken.go.jp,
             URL: http://genome.gsc.riken.go.jp/
             Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh
             ,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new
             genes. Genome Res. 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
             Matsubiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
             ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
             Hayashizaki,Y.
             Riken integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer. Genome Res.
             10 (11), 1757-1771 (2000)
             Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
             ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
             encyclopedia: real-time sequence clustering for construction of a
             nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaoka,I., Aizawa
             ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
             Hayashizaki,Y.
             Computational Analysis of Full-length Mouse cDNAs Compared with
             Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
             further details.
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
FEATURES     Location/Qualifiers
             1..649
             /organism="Mus musculus"
             /strain="C57BL/6J"
             /db_xref="taxon:10090"
             /clone="6430407824"
             /clone_lib="RIKEN full-length enriched, adult male
             olfactory bulb"
             /sex="male"
             /tissue_type="olfactory brain"
             /dev_stage="adult"
             /lab_host="DH10B"
```



```
Location/Qualifiers
1. 625
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C130099N17"
/clone_lib="RIKEN full-length enriched, 16 days embryo
```

611 AACAAATCCCATGC 624

RESULT 12	767 bp	mRNA	linear	EST 26-OCT-2001
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				

BB640911
 musculus cDNA clone A730093G08 5', mRNA sequence.
 BB640911
 GI:16476069
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 767)
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,
 M., Koyu,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ono,M., Onda,
 O.Kazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 Tagami,M., Tagawa,A., Takai,Y., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Muramatsu,M. and Hayashizaki,Y., Takeda,Y., Tanaka,T., Toya,T.,
 Riken Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC) Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
 M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 Wachihi,M., Fujisake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 S., Kawai,T., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 Hayashizaki,Y., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multichipillary sequencer. Genome Res.
 10 (11), 1571-1771 (2000)
 Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y. and Hayashizaki,Y.
 Computer-assisted methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 K., Fukui,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
 Hayashizaki,Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1..767
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A730093G08"
 /clone_lib="RIKEN full-length enriched, 7 days neonate
 cerebellum"
 /tissue_type="cerebellum"
 /dev_stage="7 days neonate"
 /lab_host="DH10b"
 /note="Site 1: Sali, Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5',

[illegible]

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 646)
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
 J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 Unpublished (1997)
 Other-SSS: RPCI-11-165H16.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
 Research Genet cs (inforesgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.

Location/Qualifiers
1. .646

Query Match	13.3%	Score 357;	DB 12;	Length 646;
Best Local Similarity	99.2%	Pred. No. 8.4e-57;		
Matches 357; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0

TITLE
JOURNAL
COMMENT

FEATURES
SOURCE

```

e mouse
tissues.
location/Qualifiers
1. 628
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="C230079M02"
/clone_lib="RIKEN full-length enriched, 0 day neonate
cerebellum"
/tissue_type="cerebellum"
/dev_stage="0 day neonate"
/lab_host="DH10b"
/notes="Site.1: SalI; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15',
GAGAGAGACAGACATCCAGAGCTCTTTTTTTTTTTTTTTVN 3', cDNA was
prepared by using triethanolamine thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA was through one round of normalization
to Rot = 20.0 and subtraction to Rot = 479.0. Second
strand cDNA was prepared with the primer adapter of
sequence 15' GAGAGAGACATCTTCAGATTAAATTAATATCCCCCCCCCCC

```


10

468 ACACCTATGGAGCAGCATTCGGGCAACACTTCCACAGATGAAGAGACCTTACCACCACTGC 409
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 468 ACACGTTAGGACGAGCATTCGGGCAACACTTCCAGATGAAGAGACTTACCAACCATCG 409


```

QY 623 ACCGAGACATCACTGGGATCTGACAAATTACCTAGCAGCCACCTTCACA 682
DB 608 ACCGTAGCTGCAAAACCCGCAATCTCTCTACCCCTGGACAGCAGGACAG 667
QY 683 AATGATTTGATGATTTTGAATTCACCAACTACACTCCCAAAAAGAGTTCAGC 742
DB 668 CTTGGGGTCTGGGCTCCATGGAATCGCCAGTAATCCATGGTCTCCATCCGGG 727
QY 743 ACTTGTAACTGACTCTGACTTACCGCCGATTCAGATATCCAGTCCCTTATGCT 802
DB 728 GCATCTCACTAGGCTGAGTATCGGCGTATTCAGATATCTTGTACCTACGGTC 787
QY 803 TCTTGACGGTAAACACAAATCCCTTCCTGTTTGAAGTCCAAAGAGAAATGCTG 862
DB 788 GCTTGAGCCTCTCTCTGGGCC--CACATCCCACTACCGGCCCAAAAGCAGATGGCTG 844
QY 863 GCTGGGTTGATTAATCTGGAACCCCTGAGCATGCGAGTCAAGTATTAATGAGCTAA 922
DB 845 CTTGGGTGATCAGCAATTTCCAGAGAGCGCAGAGCTGCAAAAGCTGTACCGGCTGG 904
QY 923 GCAAAAGCATGGAATTCATACCTAGGCGCAACATTTGAGAAATGTCAATGATAAA 982
DB 905 CCGCTCATCTGAGGTGATGCTGTTGGTGGCGCAGCGAGCGGCCCTATGGCTAAATT 964
QY 983 ATTGTATCTTACCATATCTGCTGTGTAATTTTATCTTCTTGAATAATTCATCCACA 1042
DB 965 GTCTGCGCCCACTTGGCCCGGTACCGCTTCTTACCTGAGCTTGAAGTCAACGATC 1024
QY 1043 AGGATTACATACAGGAAAGCTAT--ACAATGCTTTTCTGCTGCTCTGATCCGTTG 1099
DB 1025 GGGACTACATACATGAAATTTCTGGGCAATCCCTGGGCGGTGCTGTGATCCGCTGG 1084
QY 1100 TTTGGGACCATCTAGGGAATACTAGAAATTAATATTCAGCAGATTCATTCATG 1159
DB 1085 CGCTGGGACCTCTCGGGCCACCTACGAGAGCTTTTGTGCCACAGATGCTTGTACAG 1144
QY 1160 TGGAGATTATTAATCTCCAGTAGCTAGC 1190
DB 1145 TGGAGACTTACAGCTCTGCCGTGAAGCTGGC 1175

```

RESULT 2

```

PCT-US96-06427-1
Sequence 1, Application PC/TUS9606427
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
City: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06427
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,151
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/278WO1
TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; FAX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-06427-1

```

```

Query Match 4.7%; Score 127; DB 5; Length 1814;
Best Local Similarity 52.3%; Pred. No. 1,86-20; Mismatches 285; Indels 6; Gaps 2;
Matches 330; Conservative 0;

```

```

QY 563 GATGGCATCTGCAAGGAGCCGTTCACTGATACCAAAATCCATGCACTTGTATCATC 622
DB 548 GTTCCGCTGTCAGTCTTACCGAGCCTGCTAGCCAGTGTGATGCTGTCTTCCACC 607
QY 623 ACCGAGACATCAGTTGGGATCTGACAAATTAATCTACGAAAGTAGCCACCTTCAGA 682
DB 608 ACCGAGAGCTGCAAAACCGGCAATCTCTCTACCCCTGACACAGAGGCGACAGGACG 667
QY 683 AATGATTTGATGATTTTGAATTCACCAACTACACTCCCAAAAAGAGTTCAGC 742
DB 668 CTTGGGGTCTGGGCTCCATGGAATCGCCAGTAATCCATGGTCTCCATCCGGG 727
QY 743 ACTTGTAACTGACTCTGACTTACCGCCGATTCAGATATCCAGTCCCTTATGCT 802
DB 728 GCATCTCACTAGGCTGAGTATCGGCGTATTCAGATATCTTGTACCTACGGTC 787
QY 803 TCTTGACGGTAAACACAAATCCCTTCGTTTGAAGTCCAAAGAGAAATGCTG 862
DB 788 GCTTGAGCCTCTCTCTGGGCC--CACATCCCACTACCGGCCCAAAAGCAGATGGCTG 844
QY 863 GCTGGGTTGATTAATCTGGAACCCCTGAGCATGCGAGTCAAGTATTAATGAGCTAA 922
DB 845 CTTGGGTGATCAGCAATTTCCAGAGAGCGCAGAGCTGCAAAAGCTGTACCGGCTGG 904
QY 923 GCAAAAGCATGGAATTCATACCTAGGCGCAACATTTGAGAAATGTCAATGATAAA 982
DB 905 CCGCTCATCTGAGGTGATGCTGTTGGTGGCGCAGCGAGCGGCCCTATGGCTAAATT 964
QY 983 ATTGTATCTTACCATATCTGCTGTGTAATTTTATCTTCTTGAATAATTCATCCACA 1042
DB 965 GTCTGCGCCCACTTGGCCCGGTACCGCTTCTTACCTGAGCTTGAAGTCAACGATC 1024
QY 1043 AGGATTACATACAGGAAAGCTAT--ACAATGCTTTTCTGCTGCTCTGATCCGTTG 1099
DB 1025 GGGACTACATACATGAAATTTCTGGGCAATCCCTGGGCGGTGCTGTGATCCGCTGG 1084
QY 1100 TTTGGGACCATCTAGGGAATACTAGAAATTAATATTCAGCAGATTCATTCATG 1159
DB 1085 CGCTGGGACCTCTCGGGCCACCTACGAGAGCTTTTGTGCCACAGATGCTTGTACAG 1144
QY 1160 TGGAGATTATTAATCTCCAGTAGCTAGC 1190
DB 1145 TGGAGACTTACAGCTCTGCCGTGAAGCTGGC 1175

```

RESULT 3

```

US-07-914-281-13
Sequence 13, Application US/07914281
Patent No. 3324663
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

```


Tue Oct 8 10:18:32 2002

us-09-744-748-5.rn1

Page 4

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

```
QY 490 ATTCTGCTGGTGGTGGGCTATTTGGGACAGCTTTGACCTTACATCTGCGCAAGATG 549
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ATCTGCTGGTGGGCTATTTGGGACAGCTTTGACCTTACATCTGCGCGCAAGATG 377
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 TT---CAACATCCAGAGATCCATTCACACAGGACGCTTACTGTACCAACAAATCCAT 606
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTGCTGGGACGCTGCTGCTACATCTACGCGGACCGCAAGGTATTCACAGGACGAC 437
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 GCAGTTCTGATTCATCCAGACATCATGTTGGATCTGA---CAATTTACTTCAGCA 663
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GCGGTCTATCTGACCCAGAGATCATATACACCCAGTCCGCGGACGCTCC 497
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 664 GCTAGGCGACCTTCCAGAAATGATTTGGATGAATTCACCACTCACACTCC 723
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CCGAGGCGGACGCGGACGATGCTGTTGAGATGAGTCCCAAGCCACTGCTGG 557
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 724 CAAAAGATGGCATGTGACACTTGTAACTGACTGCTGACCTTACCGCGGTGATTCAGAT 783
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CAGCTGAAAGCATGACGATGACTTCAATCTACCATCTGCTACCGACGACCTCGAC 617
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 784 ATCCAGTGGCTTATGGCT-----TCTTGACGTTAAGCAACAATCCCTTCTGTTT 834
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 ATCTGCTGGGCTTACGCTGCTGAGCCGCTGGTCCGCGGACGCTCCGCGGACGCTC 677
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 835 GAAGTGGCAAGCAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AACCTTGGGCAAGCAAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 895 GCCAGAGTCAAGTATTAACAATGAGCTTAAGCAAAAGCAATTCATCTACCGGCA 954
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCCAGAGTGGGCTTACATCCAGAGCTTACCGGCTTACAGGCTGCTGCTGCTGCTGCT 794
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 955 GCATTTGAGAAATATGCTATGATTAATAATTTGATTCCTACATCTGCTTAAATTT 1014
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CCGTCCCAAGCAAGCCCTGCGGCGGAGGATGATGAGAGGCTGCTGCTGCTGCTGCTG 854
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1015 TATCTTCTCTTTGAAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1071
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 TATCTGCTCTGAGAACTCTTGGACCCGACATCAATCAATCAATCAATCAATCAATCAAT 914
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1072 GCTTTTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1131
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 915 GCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1132 TATATTCACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1191
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 975 TTCTGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1034
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1192 AAGTATCTGAAGAGATGCAACAAACAAATTAATTAATTAATTAATTAATTAATTAAT 1251
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 CCGTACCTGAGAGCTGAGCAAGACACCGCGCTACTGAGTACTGCTGCTGCTGCTG 1094
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1252 AAG 1254
  |||
Db 1095 GAG 1097
  |||
```

```
RESULT 5
; US-08-525-058A-13
; Sequence 13, Application US/08525058A
; Patent No. 5770420
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P. C.
; STREET: 1735 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-525-058A-13
; Query Match 4.2%; Score 113.4; DB 1; Length 1654;
; Best Local Similarity 51.2%; Pred. No. 2.5e-17;
; Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
```



```

OY 1072 GCTTTTCGGCGGCTGTGTACCTGTGTCTGTGGACATCTAGGAAAACCTATGACAT 1131
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 915 GCCCTGGAGAGGCTTGGGCGCGTGCCTGCCGTGTGCTGGGCCCCAGCAAGACACTACGAGAG 974
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1132 TATATTCACGACGATTCATTCAATTCATGTGGAAGATTATATCTCCAGTACGACAGCA 1191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 975 TTCCTGACACCCGACCGCTTCATCCAGCTGAGACACCTCCAGAGCCCCCAAGACCTGGCC 1034
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1192 AAGTATCTGGAAGGAAGTGCAGAAAACATTAAGTTATTAACCTTAAGTCTTAACTGGAG 1251
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1035 CGGATACCTGACGAGACTGAGCAAGACCAACGCCGCTACCTGAGACTCTTTCGCTGGCG 1094
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1252 AAG 1254
      ||
DB 1095 GAG 1097

RESULT 6
US-08-696-731-13
; Sequence 13, Application US/08696731
; Patent No. 5955347
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,731
; FILING DATE: 14-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/393,246
; FILING DATE:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-696-731-13

```

Query Match	4.2%	Score 113.4	DB 2	Length 1654
Best Local Similarity	51.2%	Pred. No. 2.5e-17		
Matches 401; Conservative	0	Mismatches 361	Indels 21	Gaps 5

QY	490	ATTGCGTGGTGGGCGCATTTGGGAGACGCTTTGACCTTACATCCTGGCAGCATG	549
Db	318	ATTCGCTGTGGACGTGGCCCTTTTAAACAACCCCATAGCTGTGCCCGCTGTCTCAGAGTGG	377
QY	550	TT---CAACATTCGAAGATGGCCATCTCCAAACGCGACCTTACTGTATCAACAAATTCCTAT	606
Db	378	GTGCCTGGCAGGGCGTGCATCGAACATCTACCTGCCAGCCGCAAGTGTATTCACAGCGAC	437
QY	607	GCAGTTCTGATTCATCACCGAGACATCATGTTGGATCTGA--CAAAATTACTTGACAA	663
Db	438	GGGTCATCTGTGCACACCGAGAGGTATGTACAAACCCCATGTGCCACACTCCACAGCTCC	497
QY	664	GCTATGCCACCCCTTCCAAATGATGTTGGATTTGGATGAAATTTGGATCAACCAACACACTGCC	723
Db	498	CCGAGGCGCGCAGGGGCGACGATGATGTCTGTGTACACATGAGTGTCCCAAGCAGCTGCTGG	557
QY	724	CAAAAGATGTGCATTGAGCACTTGTTTAACTGACTCTGACTTACCGCGCTGATTCAGAT	783
Db	558	CAGCTGAAGGCGATGAGCGGATCTTCAATCTCAACCATGTCTCTACCGCAGGAGCTCGAC	617
QY	784	ATCCAAAGGCTTATNGCT-----TCTTGACGGTATACCAATTCCTTGTGT	834
Db	618	ATCTTCACGCGCTACGGGTGGCTGGAGCGCTGTGTCCGGCCAGCTGCCACACCGCTC	677
QY	835	GAAGTGCACCAAGAAATTTGTTGTGGTGTGTGATGTAAGTGCACCGCTGAGCAT	894
Db	678	AACTCTGTGGCGCAAGACCGAGCTGTGGCTGTGGGAGTGTCCAACTGGGGGCAAACTCC	737
QY	895	GCCAGATGCAGATTTCAATGATGACTAAGCAAAAGCATTTGAATTCATATCTACGGCGAA	954
Db	738	GCAGAGGCGCGTACTACCAAGCTCGAGGCCCATCTCAAGGTGGAGGTGTAGCG--A	794
QY	955	GCATTTGAGAGATATGTCAATGATTAATTTGATTTCTTACCATATCTGTTGTAATTT	1014
Db	795	CGCTCCCAACAGCCCTGCCCCAGGAAACATGATGAGACGCTGTCCCGGTACAGATTC	854
QY	1015	TATCTTCTCTTGAATAATTCAATCCACAAAGATGATCATCACGGAATAT---ACAT	1071
Db	855	TATCTGGCCTTGAGAACTCTGTGACCCCGACCTCATACACGAAAGCTGTGGAGGAC	914
QY	1072	GCTTTTCTGCGCTGTGTGTACCTGTGTTTCTGGGACATCTAGGGAATATGAGCAT	1131
Db	915	GCCCTGAGGCGCTGGCGCGCTGTGCTGTGGGGCCCCACAGAACGATACGAGAGG	974
QY	1132	TATATTCAGCAGATTCATTCATTCATGTGGAAGTTATTACTCTCCAGTACGCTGCA	1191
Db	975	TTCTCTGCCACCGGAGCCCTTCTATCTACCTGTGACGACTTCCAGAGGCCCAAGAGACTGGCC	1034
QY	1192	AAGTATCTGAAGGAAGTGCACAAAAACAATTAATTAATCTTAACTTTAACTTGAGAG	1251
Db	1035	CGGTACTCTGACGAGACTGTGACAAAGACCAAGCCGCTACTGTAGGTACTTTCGCTGGCG	1094
QY	1252	AAG 1254	
Db	1095	GAG 1097	

```

: RESULT 7
: US-09-042-531-13
: Sequence 13, Application US/09042531
: Patent No. 6268193
:
: GENERAL INFORMATION:
:
: APPLICANT: LOWE, JOHN B.
:
: TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
:
: TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
:
: TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
:
: NUMBER OF SEQUENCES: 14
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: OBOLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
:
: ADDRESSEE: P.C.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
:

```

Tue Oct 8 10:18:32 2002

us-09-744-748-5.rni

Page 6

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042.531
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393.246
FILING DATE:
APPLICATION NUMBER: US 08/220.433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914.281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-13

Query Match	4.28:	Score 113.4:	DB 4:	Length 1654:
Best Local Similarity	51.28:	Pred. NO. 2.5e-17:		
Matches 401; Conservative	0;	Mismatches 361;	Indels 21;	Gaps 5

OY	490	ATTTCGGTGGGTGGGCGCATTTTGGGGCGAGCCTTTGACCTTACATCCGCGAAGCAT	549
OY	490	ATTTCGGTGGGTGGGCGCATTTTGGGGCGAGCCTTTGACCTTACATCCGCGAAGCAT	549
Db	318	ATTCGCTGTGGAGCTGGGCTTTTAAACAACCCATATGCTCTGCCCCGCTGTCAAGATG	377
OY	550	TT---CAACATCCAAAGATGCCATCTCACACGGGACCGCTTCACTGTACAACAATTCAT	606
Db	378	GGGCGTGGAGCGGTACTGACCAACATCATGTGCGACGCGGAAGGTGATCCACAGGGATAC	437
OY	607	GCAGTTGTGATCCATCACCGGAGACATCACTTGGGATCTGA---CAAAATTACTTCAGCAA	663
Db	438	GGGTCATCTGTGTACACCGACCGAGAGGCTCATGTACAAACCCGAGTCCACGCTCCACGCTCC	497
OY	664	GCTAGGCCACCCCTTCCAGAAATGATTTTGGATGAAATTGGATTCACCACTCACACTGCC	723
Db	498	CCGAGCGGCGACAGGGCAGGAGATGGATCTGTTCAAGCATGAGTCCCAAGGCCACACTGCTGG	557
OY	724	CAAAAGAGTGGGATGAGAGCACTTGTTTAACCTGCAGCTCTGATTTACCGCTGTATACAT	783
Db	558	CAGCTGAAAGCCATGACGAGGATCTTCAATCTCACCATATCTCTACCGCAGCGACTCGAC	617
OY	784	ATCCAGTGCCTTATTTGGCT-----TCTTACAGGTAAGCACAATCCCTTCGTGT	834
Db	618	ATCTTCACGCGCCCTACGGGCTGGGCTGGAGCGCGTGTGCGGGCAGCGCTCCACCGACCGGATC	677
OY	835	GAAGTGCACAAAGAGAAATTTGGTGTGCTGGTGTGTAGTAACCTGGAACCCCTGAGCAT	894
Db	678	AACCTCTGGGCCAAGACGAGCTGGGTGGCTGTGGCAGTGTCCAACTGGGGGCGCAAACTCG	737
OY	895	GCCAGAGTCAATATTAATCAATGAGCTTAACCAAAAGCATTTGAATTCATACCTACGGGCAA	954
Db	738	GCCAGAGTGCCTACTACCAAGAGCTGCAGGCGCCATCTAAGATGGAGCTGTACGG---A	794

QY	955	GCATTGGAGAAATGTGCATGTATTTAAATTTATTTCTTACCATTATCTCTGTAAATTT	1014
Db	795	CGCTCCCAAGCCCCCGCCCAAGGAGAACCATGATGAGACCTGTCCCGGTACAAAGTTC	854
QY	1015	TATCTTTCCTTTGAAATTTCAATCCCAAGAGATTACATCAGGAAAGCTAT--ACAAT	1071
Db	855	TATCTGGCCTTCGAGAACTCTCTTGCAOCCCGATTCATATCCGAGAAAGTGTGGAGAAC	914
QY	1072	GCTTTTCTGGCTGCTGTGTACTGTGTGTCTGGAGCAATCTAGGAGAAACATATGAGAT	1131
Db	915	GCCTTGGAGGCTGTGGCCGTGCCCTGTGTGTCTGTGGGCCCAAGCAAGCAACTACAGAGG	974
QY	1132	TATATTCACCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	1191
Db	975	TTTCTCGGCACCCGACGCCCTTATTCACGTGTGAGACTTCCAGAGCCCAAGSACCTGGCC	1034
QY	1192	AAGTATCTGAAGGAAAGTCGACAAACAAACATTAAGTTTATCTTACTTTCATCTGGAGG	1251
Db	1035	CGGTACTCTGCAAGAGCTGGACAAAGAACACAGCCCGCTACTGTAGCTACTTGTGCTGGCG	1094
QY	1252	AAG 1254	
Db	1095	GAG 1097	

RESULT 8
US-07-914-281-12
; Sequence 12, Application US/07914281

GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: P. C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patentin Release #1.0, Version #1.25
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/07/914,281

```

FILING DATE: 19920720

;
; CLASSIFICATION: 530
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.

```

; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-555
;

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-4500

TELEFAX: (703) 486-2347
TELEX: 248855 OPAT UR

```

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1086 base pairs
TYPE: NUCLEIC ACID

STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)
; MS-07-914-281-12

00 07 014 201 12
Query Match 4 39: 5

Query Match	51.26%	Score	112.12	DB 1	Length	1080			
Best Local Similarity	51.6%	Pred.	4.1e-17						
Matches	364	Conservative	0	Mismatches	323	Indels	18	Gaps	4

Tue Oct 8 10:18:32 2002

us-09-744-748-5.rni

Page 8

Db 868 TGCATCCAGTGGAGACTTCAGAGCCCAAGAGCCTGGCCCGGTACCTGCAGAGAGCTG 927
Oy 1210 GACAAAACATAGTATACCTTACTTACTTAACTGGAGAG 1254
Db 928 GACAAGGACACGCGCTACTGAGTACTTTCGTGGCGGAG 972

RESULT 10
US-08-525-058A-12
Sequence 12, Application US/08525058A
Patent No. 5770420
GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525.058A
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-525-058A-12

Query Match 4.2% Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 4,1e-17;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

Oy 565 TCCGATCTCACAGGAGCGTTCATCTGACAAATCCCATGCTTGTATCATCAG 624
Db 271 TCCGATCTCACAGGAGCGTTCATCTGACAAATCCCATGCTTGTATCATCAG 624
Oy 625 CGAGACATCTTGGGA--TCTGACAAATTTACTCTCAGCAAGTACCCCTTTCAG 681
Db 331 TGGGATATCTCAACCCCTTAAGTCCACCTTCCCAAGCCGAGGGGAG 390
Oy 682 AATGATTTGGATGAATTTGGATACCACTCACTCCCAAAAGAGTGGCATTTAG 741
Db 391 CGCTGATCTGTTAACTTGGAGCCACCCCTAATCTGCACACCTGGAAGCCCTGAG 450
Oy 742 CACTGTTTAACTGATCTGATACCGCGGTGATTCAGATATCAGATGAGTCTTATGAG 801
Db 451 AATATCTTAACTGATCTGATACCGCGGTGATTCAGATATCAGATGAGTCTTATGAG 801
Oy 802 TCTTGAGCGGTAAAGCAAAATCCCTTCTG-----GTTGAGTGGCAAGCAAGAG 852
Db 802 TCTTGAGCGGTAAAGCAAAATCCCTTCTG-----GTTGAGTGGCAAGCAAGAG 852

Db 511 TGGCTGAGCCGTGTCGCCAGCCTTGCCACCACCCAGCTCAACCTTCGCCAAGACC 570
Oy 853 AATTTGCTGCTGCTGTTGTTGAGTAAGTAAAGCCCTGAGCAGTCAAGTATTAAC 912
Db 571 GAGCTGGGGCTGGGGGCTGTCACACTGGAAGCGGAGCTACAGAGGCTGCTACTAC 630
Oy 913 AATGAGCTAAAGCAAGCATTTGAATTCATATCAGAGGAGAGATTTGAGAATATGTC 972
Db 631 CAGAGCCGTGAGGCTCATCTCAAGTGAAGTGAAGAGAGAGCTGCCAAGAGCCCTGAGC 690
Oy 973 AATGATTAATTTGATTTCCATCATATCTGCTTGTATTTATTTCTCTTGAAT 1032
Db 691 AAGGGAGCA--TGATGAGAGCCTGTCCCGTACAAATTTCTGCTGCTGAGAAC 747
Oy 1033 TCAATCCCAAGATTTATCATCAGGAAAGCTAT--ACAATGCTTTTCTGCTGCTG 1089
Db 748 TCTTGACACCCGCTCATATCAGGAGAGCTGTGAGGAAGCCCTGAGAGGCTGAGCC 807
Oy 1090 GTACCTGTTGTTGGAGCATCTAGGAAACTATGAGAAATTAATTTCCAGAGATTCA 1149
Db 808 GTGCCCGTGGTGTGGGCCCCAGAGCAAGCAACTACGAGAGGTTCTCCACCCGAGGCC 867
Oy 1150 TTCAATCATGTGGAAGATTTATACCTTCCAGTACGAGTACCAAGATTTCTGAAGAGTTC 1209
Db 868 TGCATCCAGTGGAGACTTCAGAGCCCAAGAGCCTGCGGTACTGTCAGAGAGCTG 927

RESULT 11
US-08-696-731-12
Sequence 12, Application US/08696731
Patent No. 595347
GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696.731
FILING DATE: 14-Aug-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/393,246

APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)486-2347

TELEX: 24885 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-696-731-12

Query Match 4.2% Score 112.2; DB 2; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 4.1e-17;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

565 TGCCATCTCACAACGACCGCTTCACTGTACAAACAAATCCCTGACGTTCTGATCCATGAC 624
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 271 TGCCACATCACTGCGACCGGAGGTGTACCCACAGCAGACAGGTGATGTCGACAC 330
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 625 CGAGACATCACTGGGA---TCTGACAAATTTACCTCAGCAGCTAGGCCACCTTCAG 681
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 331 TGGGATATCATGTCACACCTTAAGTACAGCCTCCACCTCCCGAGGCCGAGGGCAG 390
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 682 AAATGATTTGATGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 741
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 391 CGCTGATCTGTTCACTGTGAGCCACCCCTTAAGTACGACCTGACCTGGAAGCCCTGAC 450
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 742 CACTTGTTTAATCACTGACTCTGACTTACCGCCGTGATTCAGATTCAGATTCAGATTCAG 801
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 451 AGATACATTCATCTCACCATTCTCCACCGCAGCGACCTCCGACATCTTCACGCCCTAC 510
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 802 TTCTTGACGTTAAGCACAATTCCTTGT-----GTTTGAAGTCCCAAGCAAGAG 852
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 511 TGGGTGAGCGCGTGGTCCGCGACGCTCCACCCACCGCTCAACCTCTCGGCCCAAGACC 570
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 853 AAATTTGTTGCTGCTGGTGTGATGATTCAGTGAACCTGAGCATGCCAGTCCAGTCAAGTATAC 912
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 571 GAGCTGGTGGCTGGCGGTGTCACACTGGAAGCCGACTCAGCGAGGTGGCTACTAC 630
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 913 AATGACTTACGACAAAGATTTGAATTCATCTACCTACGAGGCAAGCATTTTGGAAATATATGTC 972
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 631 CAGACCTTCGACGCTCATCTCAAGTGAAGTGTACGAGCCTCCACAAAGCCCTGGCC 690
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 973 AATGATTTAAATTTGATTTCTACCATATCTGCTGTAATTTTATCTTCTTGTAAAT 1032
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 691 AAGGGACCA---TGATGAGACGCTGTCCGGTACAGTTCTACCTGGCCCTTCGAGAAG 747
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 1033 TCAATTCACAAGATTTAATATCAAGGAAAGCTAT---ACAATGCTTTCTTGGCTGGCTCT 1089
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 748 TCTTTCGACCCCGACTTACATCACGAGAAAGCTGTGAGAAAGCCCTGAGGCCCTGGGCC 807
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 1090 GTACCTGTTGTTTGGGACCATCTTACGAGAAATTTAGAAATTTATTTTCCAGACGATTTCA 1149
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 808 GTGCCCCGTGTGTGTGGGCCCCAGAGAAAGCAATTCAGAGAGTTCTCTCCACCCGACGCC 867
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 1150 TTCAATTCATGGAAGATTTAATTTCTCCAGTAGAGTACGAAAGTATTCGAAAGATTC 1209
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 868 TTAATTCACAGTGAAGATTTCAAGAGCCCAAGAGACTGTGCCCCGTGACTGAGAGAGCTG 927
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 1210 GACAAATCAATTAATTTTACCTTACTTACTTAACTTAACTGAGGAAG 1254
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 928 GACAAGACCAAGCCCGCTACCTGAGTACTTTCGCTGGCGGAG 972
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12

US-09-042-531-12

Sequence 12, Application US/09042531

Patent No. 6268193

GENERAL INFORMATION:

APPLICANT: LOME, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEIN,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/042,531
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:
 APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994
 APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEX: 24885 OPAT UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-09-042-531-12

Query Match 4.2% Score 112.2; DB 4; Length 1086;
 Best Local Similarity 51.6%; Pred. No. 4.1e-17;
 Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

565 TGCCATCTCACAACGACCGCTTCACTGTACAAACAAATCCCATGACGTTCTGATCCATGAC 624
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 271 TGCCACATCACTGCGACCGGAGGTGTACCCACAGCAGACAGGTGATGTCGACAC 330
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 625 CGAGACATCACTGGGA---TCTGACAAATTTACCTCAGCAAGCTAAGGCCACCCCTTCAG 681
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 331 TGGGATATCATGTCACACCTTAAGTACAGCCTCCACCTCCCGAGGCCGAGGGCAG 390
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 682 AAATGATTTGATGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 741
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 391 CGCTGATCTGTTCACTGTGAGCCACCCCTTAAGTACGACGACCTGGAAGCCCTGAC 450
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 742 CACTTGTTTAATCACTGACTCTGACTTACCGCGGTGATTCAGATATTCAGATTCAGATTCAG 801
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 451 AGATACATTCATCTCACCATTCTCCACCGCAGCTCCGACATCTTCAGCGCCCTACGCG 510
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 802 TTCTTGACGTTAAGCACAATTCCTTGT-----GTTTGAAGTCCCAAGCAAGAG 852
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 511 TGGGTGAGCGCGTGGTCCGCGACGCTCCACCCACCGCTCAACCTTCTCGGCCAAGACC 570
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 853 AAATTTGTTGCTGCTGGTGTGATGATTCAGTGAACCTTGAAGTCCAGAGTCAAGTATTTAC 912
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 571 GAGCTGGTGGCTGGCGGTGTCCAATGGAAGCCGACTCAGCGAGGTGGCTACTAC 630
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 913 AATGACTTACGACAAAGATTTGAATTCATCTACCTTACGAGGCAAGCATTTTGGAAATATATGTC 972
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 631 CAGACCTTCGACGCTCATCTCAAGTGTGAGAGTGTACGAGACGCTCCACAAAGCCCTGGCC 690
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```

APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ANTI-SENSE: NO
US-07-914-281-1

```

Query Match 4.2%; Score 112.2; DB 1; Length 2043;

Best Local Similarity 51.6%; Pred. No. 5.3e-17;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

565 TGCACATCTCAACAGGACGCTTCACTGACACAAATCCATGCGAGTTCGATCCATCAC
343 TGCACATCTCACTGCGCCGACGAGGTGACCCACAGGACGAGGTCATGTCACAC
625 CGAGACATCAAGTTGGGA---TCTGACAAATTTACTGACAGAGTCCAGCCCTTCAG
403 TGGGATTCATGTCACCAACCTTAAGTTCAGCGCTCCACCTTCCCGAGCGGACGAG
682 AAATGATTTGATGATTTGATGATCACTCACTCCCAAAAGAGTGGATTGAG
463 CGCTGATCTGCTTCACTTGGAGCCACCCCTTACTGACGACGACCTGGAGCCCTGGAC
742 CACTGTTTAACTGACTCTGACTTACCGCGGTGATTCAGATATCCAAGTCCCTATGGC
523 AGATACTTCAATCTACCATCTCTCCGAGGACCTCCACATCTTACAGCCCTTACGGC
802 TTCTTGACGGTAAGCAAAATCCCTTCGT-----GTTTGAAGTCCCAAGCAAGAG
583 TGGCTGAGCGGTGCTGCTCCGACGCTGCCACCCACCGCTCAACCTCTGCGCCAAAGC
853 AAATGCTGTGCTGCTGCTGAGTAACTGGAACCTGAGATGCGACAGTCAAGTATTAC
643 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
913 AATGACCTAAGCAAAAGCAATTTGAATCAATCACTACGCGGCAAGCATTTGAGAAATATG
703 CAGAGCCTGAGGCTCATCTCAAGGTGAGGTGAGGAGGTCCTCCCAAGCCCTGCCC
973 AATGATTAATTAATTTGATTTCTTACCATATCTGCTTGAATTTATCTTCTTGAAT
763 AAGGGAGCA---TGATGAGAGCGCTGCTCCGATCAAGTTTCACTGCGCTGCAAGC
1033 TCAATCACAAGATTCATCACTGAGAAAGCTAT---ACAATGCTTTTCTGCTGCTGCT
820 TCTGTGACCCCGCATCTACATCACCAGAAAGCTGTGAGGAAAGCCCTGAGAGCCTGGCC
1090 GTACCTGTTTGTGGGACCATCTAGGAAAGCACTATGAGATTAATATATTCAGCAGATTCA
880 GTGCGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1150 TTGATTTATTTGAGAAATTTAACTCTCCAGAGAGTACGAAAGTTCGAGAGAGATC
940 TTGATTCACATGAGAGCTTCAAGAGCCCAAGGACCTGCGCTGACCTGCAAGAGAGCTG
1210 GACAAACATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
1000 GACAAAGACCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

```

RESULT 15

US-08-393-246-1

; Sequence 1, Application US/08393246

; Patent No. 5595900

GENERAL INFORMATION:

APPLICANT: LONE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

ADDRESS: P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/393,246

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/220,433

FILING DATE: 30-MAR-1994

APPLICATION NUMBER: US 07/914,281

FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2043 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ANTI-SENSE: NO

US-08-393-246-1

Query Match 4.2%; Score 112.2; DB 1; Length 2043;

Best Local Similarity 51.6%; Pred. No. 5.3e-17;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

565 TGCACATCTCAACAGGACGCTTCACTGACACAAATCCATGCGAGTTCGATCCATCAC
343 TGCACATCTCACTGCGCCGACGAGGTGACCCACAGGACGAGGTCATGTCACAC
625 CGAGACATCAAGTTGGGA---TCTGACAAATTTACTGACAGAGTCCAGCCCTTCAG
403 TGGGATTCATGTCACCAACCTTAAGTTCAGCGCTCCACCTTCCCGAGCGGACGAG
682 AAATGATTTGATGATTTGATGATCACTCACTCCCAAAAGAGTGGATTGAG
463 CGCTGATCTGCTTCACTTGGAGCCACCCCTTACTGACGACGACCTGGAGCCCTGGAC
742 CACTGTTTAACTGACTCTGACTTACCGCGGTGATTCAGATATCCAAGTCCCTATGGC
523 AGATACTTCAATCTACCATCTCTCCGAGGACCTCCACATCTTACAGCCCTTACGGC
802 TTCTTGACGGTAAGCAAAATCCCTTCGT-----GTTTGAAGTCCCAAGCAAGAG

```

```
Db 583 TGCGTGGAGCCGCTGTCGCCGCCAGCCTGCCCACCCACCGCTCAACCTCTGGGCCAAGACC 642
OY 853 AAATTGGTGTCTGGTGTGTGACTAACTGGAACCTGAGCAGCCAGAGTCAAGTATTAC 912
Db 643 GAGCTGTGGCTGGGCGGTGTCCAACTGMAAGCCGACTCAAGCCAGGGTGGCTACTAC 702
OY 913 AATGAGCTAAGCAAAAAGCATTTGAATCCATACCTACGGGCAAGCATTTGAGAAATATGTC 972
Db 703 CAGAGCCTTCAGGCTCATCTCAAGGTGAGAGTGTACGAGCTGCCACAAGCCCCTGCCC 762
OY 973 AATGATAAAAATTTGATTCCTACCATATCTGTTAAATTTATCTTTCTTTGAAAAAT 1032
Db 763 AAGGGGACCA---TGATGAGAGCGCTGTCCGGTACAAAGTTCTACCTGGGCTTCGAGAAC 819
OY 1033 TCATCCACAAAGGATTAATCATCAGGAAAAAGCTAT---ACAATGCTTTCTGGCTGGCTCT 1089
Db 820 TCCTTGACCCCCGACTACATACACCGAAGAGCTGTGAGAGAAAGCCCTGGAGGCGCTGGGCC 879
OY 1090 GTACCTGTGTCTGGGACCATCTAGGAAAACTATGAGAAATTAATTCACAGCAATCA 1149
Db 880 GTGCCCGTGTGCTGGGCCCCACAGAGCACTACGAGAGGTTCTGCGACCCGACGCC 939
OY 1150 TTGATTCATGTGGAAGATTAATTAATCTCTCCAGTGAAGTACCAAGTATCTGAAGAAATC 1209
Db 940 TTGATTCAGCTGAGACGACTTCCAGAGCCCAAGAGACTGGCCCGGTACCTGACAGAGCTG 999
OY 1210 GACAAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1254
Db 1000 GACAAAGACCAAGCCCGCTACTGAGCTACTTTCGCTGGCGGGAG 1044
```

Search completed: October 6, 2002, 07:21:51
Job time : 68.6005 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:39:45 ; Search time 1686.68 seconds
(without alignments)
16699.717 Million cell updates/sec

Title: US-09-744-748-5_COPY_29_1374

Perfect score: 1346
Sequence: 1 accgctctccctcgagctc.....agaatggtttggaattaa 1346

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :
1: gb_da : *
2: gb_hcg : *
3: gb_in : *
4: gb_om : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vl : *
15: em_ba : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_mu : *
20: em_om : *
21: em_or : *
22: em_ov : *
23: em_pat : *
24: em_ph : *
25: em_pl : *
26: em_ro : *
27: em_sts : *
28: em_un : *
29: em_vl : *
30: em_hcg_hum : *
31: em_hcg_inv : *
32: em_hcg_other : *
33: em_htgo_inv : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1346	100.0	3019	9	AB023021	AB023021 Homo sapi
2	1171.2	87.0	2501	9	HS238701	AJ238701 Homo sapi
3	1086.4	80.7	80247	9	AL512406	AL512406 Human DNA
4	1009.6	75.0	2156	10	AB049819	AB049819 Rattus no
5	986.8	73.3	2139	10	AB015426	AB015426 Mus muscu
6	958.2	71.2	1128	10	AF345993	AF345993 Rattus no
7	933.6	69.4	1705	10	AF230460	AF230460 Cricetulu
8	789	58.6	1707	5	AB035906	AB035906 Gallus ga
9	676.8	50.3	1080	5	AB035905	AB035905 Xenopus 1
10	639	47.5	71396	2	AC100674	AC100674 Mus muscu
11	277.8	20.6	1567	5	AB023627	AB023627 Danio rer
12	234.6	17.4	1832	5	AB035628	AB035628 Xenopus 1
13	222.6	16.5	1092	5	AB035907	AF288369 Gallus ga
14	220.2	16.4	1264	5	AF288369	U73678 Gallus gall
15	176.2	13.1	2657	5	GGU73678	AL034348 Human DNA
16	169.4	12.6	120514	9	HS84D21	AL356303 Homo sapi
17	169.4	12.6	191190	2	AL356303	BC001879 Homo sapi
18	169	12.6	1180	9	BC001879	U78737 Cricetulu
19	145.2	10.8	1429	10	CGU78737	AF090450 Cricetulu
20	145.2	10.8	1462	10	AF090450	AF090449 Cricetulu
21	143.6	10.7	1718	10	AF090449	AB035908 Gallus ga
22	137.2	10.2	1044	5	AB035908	X87810 B. taurus al
23	132	9.8	2293	4	BTA134FTG	AJ132772 Bos tauru
24	132	9.8	12607	4	BTA132772	AC100674 Mus muscu
25	128	9.5	71396	2	AC100674	AB039126 Mus muscu
26	127.6	9.5	1158	10	AB039126	AF016899 Schistos
27	127	9.4	1487	3	AF016899	AF028686 Sequence
28	127	9.4	1814	6	AR028686	U45980 Mus muscu
29	127	9.4	3594	10	MMU45980	AC091284 Mus muscu
30	127	9.4	117993	2	AC091284	AJ132776 Bos tauru
31	125.6	9.3	1613	4	BTA132776	AJ132775 Bos tauru
32	125.6	9.3	1619	4	BTA132775	AJ132774 Bos tauru
33	125.6	9.3	1644	4	BTA132774	AJ132773 Bos tauru
34	125.6	9.3	1873	4	BTA132773	AF345881 Macaca mu
35	124.6	9.3	1119	9	AF345881	AB039127 Mus muscu
36	122.8	9.1	1170	10	AB039127	AB039131 Mus muscu
37	122.8	9.1	1170	10	AB039131	U33457 Mus muscu
38	122.8	9.1	1615	10	MMU33457	AF345884 Saliniri s
39	122	9.1	1113	9	AF345884	AB039124 Mus muscu
40	121.2	9.0	1170	10	AB039124	AB039125 Mus muscu
41	121.2	9.0	1170	10	AB039125	AB039128 Mus muscu
42	121.2	9.0	1170	10	AB039128	AB039129 Mus muscu
43	121.2	9.0	1170	10	AB039129	AB039130 Mus muscu
44	121.2	9.0	1170	10	AB039130	AB039132 Mus muscu
45	121.2	9.0	1170	10	AB039132	

ALIGNMENTS

RESULT 1
AB023021
LOCUS
DEFINITION
AB023021 Homo sapiens FUT9 mRNA for alpha-1,3-fucosyltransferase IX, complete cds.
ACCESSION
AB023021 GI:5139692
VERSION
AB023021.1
KEYWORDS
alpha-1,3-fucosyltransferase IX, FUT9.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 (sites)
Nakagawa, S., Sasaki, K., Shilina, T., Inoko, H., Saitou, N. and Narimatsu, H.
Narimatsu, H.
Alpha 1,3-fucosyltransferase IX (Fuc-TIX) is very highly conserved between human and mouse; molecular cloning, characterization and tissue distribution of human Fuc-TIX
FEBS Lett. 453, 237-242 (1999)
JOURNAL
2 (bases 1 to 3019)
REFERENCE
Kaneko, M., Kudo, T. and Narimatsu, H.
AUTHORS
TITLE
Direct Submission

FEATURES

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
295. .1374
/gene="FUT9"
295. .1374

```

Query Mat

1
2

Db 29 ACTGCTCTCCCTGACAGCTCCCGCGCCGCGCTGTGCGTCCGGTGTGCCAG 60
 Qy 61 CCCAGTCGCGCTCTTAGAGACAGCGCGCACCGCGCCGCGCTGTGCTGCTGTGCCAG 88
 Db 89 CCCAGTCGCGCTTAGAGACAGCGCGCACCGCGCCGCTGTGCTGCTGTGCCAG 120
 Qy 121 CCGCGCAGCCCTTGGCAGCGCCCGGATGGCGCTTACCCCTAGACCGATTTAGAATGT 148
 Db 149 CCGCGCAGCCCTTGGCAGCGCCCGGATGGCGCTTACCCCTAGACCGATTTAGAATGT 180
 Qy 181 AATTAAGTCAAGATTTGATTAATACAGTGAAGTAAACAAGTGTACGTGCTGCCA 208
 Db 209 AATTAAGTCAAGATTTGATTAATACAGTGAAGTAAACAAGTGTACGTGCTGCCA 240
 Qy 241 TGAATGTCTCTATATATGAAGAAATTTAGATCAATCAATCCAAAGAAATCTTGCCAT 268
 Db 269 TGAATGTCTCTATATATGAAGAAATTTAGATCAATCAATCCAAAGAAATCTTGCCAT 300
 Qy 301 TTTTAATGTGTGATTAATCTGGCGCTTTCAATGAGCATCTCTCTAATTAACATCAAC 328
 Db 329 TTTTAATGTGTGATTAATCTGGCGCTTTCAATGAGCATCTCTCTAATTAACATCAAC 360
 Qy 361 CTACCAACAGCTGGATTTGATGTCACATGSAATGACGACGCTGTGCTGAAGAAATGAAA 388
 Db 389 CTACCAACAGCTGGATTTGATGTCACATGSAATGACGACGCTGTGCTGAAGAAATGAAA 420
 Qy 421 ACTTCTTTTCCACAAACTGATTAATTAAGAACTCTATTCTGGCTGTGGGTGTGC 448
 Db 449 ACTTCTTTTCCACAAACTGATTAATTAAGAACTCTATTCTGGCTGTGGGTGTGC 480
 Qy 481 CATTTGGGAGACCTTTGACCTTACATCTGGCAAGCATGTTGAACATCAAGATGCC 508
 Db 509 CATTTGGGAGACCTTTGACCTTACATCTGGCAAGCATGTTGAACATCAAGATGCC 540
 Qy 541 ATCTCAACAGGACGTTCACTGTACAACAATCCATGCAAGTTCTGATCCATCACCAG 568
 Db 569 ATCTCAACAGGACGTTCACTGTACAACAATCCATGCAAGTTCTGATCCATCACCAG 600
 Qy 601 ACATCAGTTGGATCTACAATTTACTCTAGCAAGTACGCCACCTTCACAAGATGA 660

ORGANISM

REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE	REFERENCE AUTHORS	TITLE	JOURNAL
Mammalia; Euteria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Hominoidei; Homo.	1 (bases 1 to 2501)		Mennesson, B., Orlol, R. and Mollicoone, R.	FYR4 and FYR9 genes are expressed early in human embryogenesis	20386591
Callitrua thomasi, A., Coullin, P., Candellier, J. J., Balanzino, L.,	2 (bases 1 to 2501)		Mollicoone, R.	Submitted	2000
Submitted	2000		Submitted	2000	

COMMENT related sequence: AB015426.
FEATURES
source location/qualifiers
1..2501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/tissue_type="whole embryo"
/dev_stage="40 to 60 days"
105..2466
/gene="FUT9"
105..1184
/gene="FUT9"
/codon_start=1
/product="alpha-3-fucosyltransferase"
/protein_id="CAB41890.1"
/db_xref="GI:4741567"
/db_xref="SPTREMBL:O9Y231"
/translation="MTSTSKILRPFLVCIIGCFMNCILYIKPTNSWTFSPWESA
SSVLKMKPFSTKTDYFNETITLVVWVFFGOTFDTSCQAFNITDGLHTDRLSYNK
SHAVLIHRDISMDLTNLPQARPFPQWIMNNLESPTHTQSKSIEHLNLTLYRR
DSDIQVPGFLVSTNPPEVPSKELVCWVSNMNEHARVKVKNLSKIEHTY
GOAREGVNDKNIPTISACKFVLSFENSINKDYITEKLYNAFLAGSVVPLGPREN
YENIIPADSFIVHEDYNSPSELAKYLKFKVDKNNKLYLSYFMWRKDFTVNLDPWESHA
CIACDHVRHDEYKSVGNLEKFMFN"
polyA_site
2192
/gene="FUT9"
/evidence=experimental
polyA_site
2235
/gene="FUT9"
/evidence=experimental
polyA_site
2310
/gene="FUT9"
/evidence=experimental
polyA_site
2466
/gene="FUT9"
/evidence=experimental
BASE COUNT 818 a 436 c 422 g 825 t
ORIGIN
Query Match 87.0%; Score 1171.2; DB 9; Length 2501;
Best Local Similarity 99.7%; Pred. No. 2.4e-286;
Matches 1173; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 429 CATCACCGAGACATACGCTGGATCTGACTTAACCTACACAGAACTAGGCCACCCTTC 488
QY 651 CAGAAATGATTTGGATGATGATTTGGAATCACCACACTCACCTCCCCAAAAGAGTGGCATT 710
DB 489 GAGAAATGATTTGGATGATGATTTGGAATCACCACACTCACCTCCCCAAAAGAGTGGCATT 548
QY 711 GAGCACTTGTTAACCTGACTGCTGACTTAACGCGCGTGAATCAGATATCCAGTGCCTTAT 770
DB 549 GAGCACTTGTTAACCTGACTGCTGACTTAACGCGCGTGAATCAGATATCCAGTGCCTTAT 608
QY 771 GCGTCTTGAGCGTAAGCAACAATCCCTGCGTGTGTAAGTGGCCAAAGCAAGAAATTTG 830
DB 609 GCGTCTTGAGCGTAAGCAACAATCCCTGCGTGTGTAAGTGGCCAAAGCAAGAAATTTG 668
QY 831 GTGTCGCGGTTGTGAGTAAGTGAACCTGAGACATGCGACAGTCAAGTATTAACATGAG 890
DB 669 GTGTCGCGGTTGTGAGTAAGTGAACCTGAGACATGCGACAGTCAAGTATTAACATGAG 728
QY 891 CTAAAGCAAAAGCATTTGAATCCATACCTACCTACGCGCAAGCATTTGAGAAATATGCAATGAT 950
DB 729 CTAAAGCAAAAGCATTTGAATCCATACCTACCTACGCGCAAGCATTTGAGAAATATGCAATGAT 788
QY 951 AAAAATTTGATTCCTACCATATCTGCTGTAATTTATCTTCCCTTGAATAATCCAATC 1010
DB 789 AAAAATTTGATTCCTACCATATCTGCTGTAATTTATCTTCCCTTGAATAATCCAATC 848
QY 1011 CACAAGGATTACATCAGCAAGAAAGCATATCAATGCTTTCTGGCGCTCTGACTGCTT 1070
DB 849 CACAAGGATTACATCAGCAAGAAAGCATATCAATGCTTTCTGGCGCTCTGACTGCTT 908
QY 1071 GTTCTGGGACCATCTAGGGAAGAACTATGAAATTTATTTCCAGCAGATTCATTCTAT 1130
DB 909 GTTCTGGGACCATCTAGGGAAGAACTATGAAATTTATTTCCAGCAGATTCATTCTAT 968
QY 1131 GTGGAAGATTATACCTCTCCAGTACGAGTACGAAAGTATCTGAAGAGTCCACAAAAC 1190
DB 969 GTGGAAGATTATACCTCTCCAGTACGAGTACGAAAGTATCTGAAGAGTCCACAAAAC 1028
QY 1191 AATAAGTTATACCTTACTTACTTAACTGGAGAAAGATTCTACTGTAAATCTTCACAGA 1250
DB 1029 AATAAGTTATACCTTACTTACTTAACTGGAGAAAGATTCTACTGTAAATCTTCACAGA 1088
QY 1251 TTTTGGGAATCACATGATGTTGGCTTGCATCATGTGAAAAAGCATCAGAATATAG 1310
DB 1089 TTTTGGGAATCACATGATGTTGGCTTGCATCATGTGAAAAAGCATCAGAATATAG 1148
QY 1311 TCTGTGTGTAATTTAGAGAAATGCTTTTGGAAATTA 1346
DB 1149 TCTGTGTGTAATTTAGAGAAATGCTTTTGGAAATTA 1184
RESULT 3
AL512406 80247 bp DNA linear PRI 27-APR-2001
LOCUS AL512406
DEFINITION Human DNA sequence from clone Rpl1-504J9 on chromosome 6, complete
sequence.
ACCESSION AL512406
VERSION AL512406.14 GI:13897154
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 80247)
Bates K.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Apr 30, 2001 this sequence version replaced gi:13396709.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Ss: SwissProt, Tr: TrEMBL, Wp: WormPeP, Information on the WormPeP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/WormPeP This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>
<http://www.sanger.ac.uk/HGP/Chr6>
 of Pieter de Jong. For further details see

<http://www.chori.org/dacpac/home.htm>

VECTOR: PBACE3.6
 IMPORTANT: This sequence is not the entire insert of clone RP11-504J9. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP3-381A2 is at 80148 in this sequence. The true right end of clone RP11-77B15 is at 100 in this sequence.

FEATURES

source
 1..80247
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP11-504J9"
 /clone_11b="RP11-11.2"
 184..1422
 /note="L2 repeat: matches 875..2419 of consensus"
 1744..1834
 /note="MIR repeat: matches 166..258 of consensus"
 2950..2977
 /note="14 copies 2 mer ca 100% conserved"
 3411..3669
 /note="LTR16C repeat: matches 93..364 of consensus"
 4372..4704
 /note="LTR1A1 repeat: matches 5..365 of consensus"
 /note="complement(7821..8406)"
 /note="match: GSS: Em:A0550748"
 8126..8184
 /note="MER58B repeat: matches 272..332 of consensus"
 8185..8696
 /note="MIR2D repeat: matches 1..513 of consensus"
 8779..8817
 /note="MIR2D repeat: matches 504..542 of consensus"
 8819..8974
 /note="MER58A repeat: matches 6..162 of consensus"
 /note="complement(9443..9891)"
 /note="match: GSS: Em:B90633"
 /note="complement(9452..9883)"
 /note="match: GSS: Em:A0119695"
 9782..9894
 /note="MIR repeat: matches 92..205 of consensus"
 10089..10212
 /note="MIR repeat: matches 135..262 of consensus"
 /note="complement(10093..10671)"
 /note="match: GSS: Em:A0237061"
 10223..10282
 /note="MIR repeat: matches 203..262 of consensus"
 /note="complement(10315..10631)"
 /note="match: GSS: Em:A0200259"
 10805..11186
 /note="match: GSS: Em:A0072353"
 11657..11844
 /note="L2 repeat: matches 2346..2750 of consensus"

misc-feature
 complement(11659..12084)
 /note="match: GSS: Em:A0518361"
 /note="complement(11674..11987)"
 /note="match: GSS: Em:A0094678"
 12083..12588
 /note="match: GSS: Em:A0691235"
 12596..12810
 /note="MIR repeat: matches 23..248 of consensus"
 12920..12966
 /note="LIM5 repeat: matches 6228..6295 of consensus"
 12987..13169
 /note="LIM5 repeat: matches 5852..6043 of consensus"
 13168..13280
 /note="LIM48 repeat: matches 5710..5832 of consensus"
 13851..17218
 /note="LIP3 repeat: matches 2787..6155 of consensus"
 18903..18953
 /note="MIR2CB repeat: matches 439..501 of consensus"
 18984..19433
 /note="MIR2CB repeat: matches 1..460 of consensus"
 /note="complement(19108..19690)"
 /note="match: GSS: Em:A0563935"
 20683..20801
 /note="AluYo/FAM repeat: matches 169..288 of consensus"
 20802..20862
 /note="AluY/FAM repeat: matches 3..66 of consensus"
 20912..21033
 /note="MIR repeat: matches 114..237 of consensus"
 21047..21128
 /note="41 copies 2 mer ta 63% conserved"
 21676..21859
 /note="L2 repeat: matches 2130..2308 of consensus"
 21890..23612
 /note="LIP repeat: matches 841..2570 of consensus"
 23601..26609
 /note="LIP2 repeat: matches 3107..6145 of consensus"
 26620..26789
 /note="MER5A repeat: matches 1..174 of consensus"
 /note="complement(28627..28837)"
 28521..28627
 /note="MER39 repeat: matches 5..111 of consensus"
 28656..28837
 /note="MER39 repeat: matches 54..235 of consensus"
 28838..29142
 /note="AluSc repeat: matches 4..308 of consensus"
 29143..29494
 /note="MER39b repeat: matches 235..549 of consensus"
 30529..31344
 /note="LIP3 repeat: matches 5331..6146 of consensus"
 32344..32340
 /note="L2 repeat: matches 2627..2750 of consensus"
 32399..32578
 /note="MIR repeat: matches 12..203 of consensus"
 32816..32979
 /note="MIR repeat: matches 91..259 of consensus"
 /note="complement(33006..33450)"
 /note="match: GSS: Em:A0261227"
 /note="complement(33058..33444)"
 /note="match: GSS: Em:A0040453"
 33312..33596
 /note="LIM repeat: matches 3473..3773 of consensus"
 34856..34915
 /note="30 copies 2 mer aa 68% conserved"
 /note="complement(34987..35547)"
 /note="match: GSS: Em:A0321598"
 match: STS: Em:G55571"
 /note="complement(36488..36936)"
 /note="match: GSS: Em:B41073"
 /note="complement(36654..37090)"
 /note="match: GSS: Em:A0772584"
 37102..37591
 /note="match: GSS: Em:A0489567"
 /note="complement(37581..38221)"
 /note="match: GSS: Em:A0418233"

repeat_region 37720..37772 /note="L2 repeat: matches 2657. .2710 of consensus"
misc_feature /note="match: GSS: Em: A0819773"
misc_feature 38218..38732 /note="match: GSS: Em: A0599868"
repeat_region 38861..39164 /note="AluX repeat: matches 1. .305 of consensus"
repeat_region 39612..39653 /note="21 copies 2 mer ac 90% conserved"
repeat_region 40854..40909 /note="L2 repeat: matches 2655. .2710 of consensus"
repeat_region 42169..42270 /note="MER5A repeat: matches 1. .111 of consensus"
repeat_region 42411..42548 /note="MER5B repeat: matches 21. .181 of consensus"
repeat_region 42780..42938 /note="MER5B repeat: matches 17. .177 of consensus"
repeat_region 43229..43326 /note="MIR repeat: matches 40. .148 of consensus"
repeat_region 43583..45428 /note="MER5B repeat: matches 18. .62 of consensus"
misc_feature /note="match: GSS: Em: A0269059"
misc_feature /note="match: GSS: Em: A0336527"
misc_feature /note="match: GSS: Em: A0801834"
repeat_region 48597..48803

Query Match 80.7%; Score 1086.4; DB 9; Length 80247;
Best Local Similarity 99.9%; Pred. No. 2.2e-264;
Matches 1087; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 259 GAAAAATTATGACATCAATCCCAAGAAATTCCTGCCCATTTTAAATGTCGATTA 318
Db 35551 GAAAAATTATGACATCAATCCCAAGAAATTCCTGCCCATTTTAAATGTCGATTA 35610

QY 319 TCCTGGGCTTTTCATGCGATGTCCTCATTTACATCAACCTACCAACAGCTGATCT 378
Db 35611 TCCTGGGCTTTTCATGCGATGTCCTCATTTACATCAACCTACCAACAGCTGATCT 35670

QY 379 TCAGTCCAAATGGAATCGCAGCTCTGCTGAAATGAATAAATCTTTCCACCAAAA 438
Db 35671 TCAGTCCAAATGGAATCGCAGCTCTGCTGAAATGAATAAATCTTTCCACCAAAA 35730

QY 439 CTGATTTATTTAAATGAATCTATTTCTGTTGGTGTGGCCATTTGGGACAGCTTTG 498
Db 35731 CTGATTTATTTAAATGAATCTATTTCTGTTGGTGTGGCCATTTGGGACAGCTTTG 35790

QY 499 ACCTTACATCTCGGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAACGACCGTT 558
Db 35791 ACCTTACATCTCGGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAACGACCGTT 35850

QY 559 CACGTCAACAATCCCAATGCAATGTCATCCACGACATCAAGTATGGAGATCTGA 618
Db 35851 CACGTCAACAATCCCAATGCAATGTCATCCACGACATCAAGTATGGAGATCTGA 35910

QY 619 CAAATTTACCTCAAGCAAGTACGACCTTCCAGAAATGATTTGGATTAATTTGGAAT 678
Db 35911 CAAATTTACCTCAAGCAAGTACGACCTTCCAGAAATGATTTGGATTAATTTGGAAT 35970

QY 679 CACCAACTCACTCCCAAAAAGATGCGATTTGACACTGTTTAACTGACTGACTT 738
Db 35971 CACCAACTCACTCCCAAAAAGATGCGATTTGACACTGTTTAACTGACTGACTT 36030

QY 739 ACCGCGCTGATTCAGATATCCAGTGCCTTATGCTCTGAGAGTAAAGCAAAATCCCT 798
Db 36031 ACCGCGCTGATTCAGATATCCAGTGCCTTATGCTCTTTCAGAGTAAAGCAAAATCCCT 36090

Db 36031 ACCGCGCTGATTCAGATATCCAGTGCCTTATGCTCTTTCAGAGTAAAGCAAAATCCCT 36090

QY 799 TCCTGTTTGAAGTGCACCAAGCAAGAAATTTGCTGCTGGTGTGAGTAACTGCAACC 858
Db 36091 TCCTGTTTGAAGTGCACCAAGCAAGAAATTTGCTGCTGGTGTGAGTAACTGCAACC 36150

QY 859 CTGAGCATGCCAGATCAAGTATTTACATGAGCTTAACCAAGCAATTTGAATCCATACCT 918
Db 36151 CTGAGCATGCCAGATCAAGTATTTACATGAGCTTAACCAAGCAATTTGAATCCATACCT 36210

QY 919 ACGGCAAGCATTTGGAGATATGTCATATGATAAATTTGATTCACCATATCTGCTT 978
Db 36211 ACGGCAAGCATTTGGAGATATGTCATATGATAAATTTGATTCACCATATCTGCTT 36270

QY 979 GTAAATTTATCTTCTTCCTTTGAAATTAATCAATCCCAAGATTAATCAATCCCAAGATTA 1038
Db 36271 GTAAATTTATCTTCTTCCTTTGAAATTAATCAATCCCAAGATTAATCAATCCCAAGATTA 36330

QY 1039 ACAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098
Db 36331 ACAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36390

QY 1099 AGAATTTATTTCCAGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1158
Db 36391 AGAATTTATTTCCAGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 36450

QY 1159 TACCAAAATATTCGAAGAGTGCACCAAAACATAAATTAATTAATTAATTAATTAATTAATTAAT 1218
Db 36451 TACCAAAATATTCGAAGAGTGCACCAAAACATAAATTAATTAATTAATTAATTAATTAATTAAT 36510

QY 1219 GGAAGAAAGATTCATCTGTAATCTTCCACGATTTTGGGAATCATGATGTTGGCTT 1278
Db 36511 GGAAGAAAGATTCATCTGTAATCTTCCACGATTTTGGGAATCATGATGTTGGCTT 36570

QY 1279 GCGATCTATGGAAGGCAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1338
Db 36571 GCGATCTATGGAAGGCAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 36630

QY 1339 GGAATTAA 1346
Db 36631 GGAATTAA 36638

RESULT 4
AB049819
LOCUS
DEFINITION
AB049819 2156 bp mRNA linear ROD 18-JAN-2002
complete cds.
VERSION
AB049819.1 GI:13591588
KEYWORDS
SOURCE
Rattus norvegicus cdna to mRNA.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (sites)
Shimoda,Y., Tajima,Y., Osanai,T., Katsune,A., Kohara,M., Kudo,T.,
Narimatsu,H., Takashima,N., Ishii,Y., Nakamura,S., Osunli,N. and
Sanai,Y.
Pax6 Controls the Expression of Lewis x Epitope in the Embryonic
Forebrain by Regulating alpha 1,3-fucosyltransferase IX Expression
J. Biol. Chem. 277 (3), 2033-2039 (2002)
11675393
PUBMED
AUTHORS
Sanai,Y
TITLE
Direct Submission
SUBMITTED (11-Oct-2000) Yutaka Sanai, Tokyo Metropolitan Institute
of Medical Science, Department of Biochemical Cell Research;
Honkomadome 3-18-22, Bunkyo-ku, Tokyo 113-8613, Japan
(E-mail:sanai@rinsoken.or.jp, Tel:81-3-3823-2101(ex.5233),
Fax:81-3-3828-6663)
FEATURES
Location/Qualifiers

[illegible]

OY	745	GATGATTCAATATATCCAAAGCGCTTATAGCTTCCTTGACGGTAAGACAATAATCCTTCGTGT	804
Db	824	GTGATTCCAAATATACAAAGGCTTTACGGCTTCTTACAGGTGAAGACCCTTTTGTGT	883
OY	805	TTCGAATGCCAAGCAAAGAATTTGGTGTGGGTTGTGAGTAACATGGAACCTGTAGC	864
Db	884	TTCGAATGCCAAGCAAAGAAGATTAGTGTGGTGTGTAGTAACATGGAATCCCGAGC	943
OY	865	ATGCCAGACTCATATATTACATAGAGCTTAAGCAAAAGCATTTGAATTCATACATGAGGC	924
Db	944	ATGCCAGAGCTCATATATTACATAGAGCTTAAGCAAAAGCATTTGAATTCATACATGAGGC	1003
OY	925	AAGCATTTGGAGAAATATGTCAATGATATAAAATTTGATCTCTACCATATATCTGCTGTAAAT	984
Db	1004	AAGCATTTGGAGAAATATGTGAATGATATAAAATCTGATTTCCACCATATATCTACTGTAAAT	1063
OY	985	TTTTATCTTCTCTTGGAAAAATTCATATCCACAGAGATTATACATACAGGAAAGGCTATACAAATG	1044
Db	1064	TTTTATCTTCTCTTGGAAAAATTCATATCCACAGAGATTATACATACAGGAAAGGCTCTACAAATG	1123
OY	1045	CTTTTCTGTGGCGCTGTCTTACCTGTGTGTCTTGGAGCATCATAGGAAACATATGAGAAAT	1104
Db	1124	CATTTTGTGGCGCTGTCTTACCTGTGTGTCTTGGAGCATCATAGGAAACATATGAGAAAT	1183
OY	1105	ATATTCACGACGATTCATCATCATATGGAAGATTTTAACTCTGCCATGAGCTAGCA	1164
Db	1184	ATATTCACGACGATTCATCATCATATGGAAGATTTTAACTCACACAGTAGAGTTGGCAA	1243
OY	1165	AGTATCTGAAGAGAGTGCAGCAAAAACAATTAAGTATACCTTTAGTTACTTTAACGTGAGGA	1224
Db	1244	AATATCTGAAGAGAGTGCAGCAAAAACAATTAAGTATACCTTTAGTTACTTTAACGTGAGGA	1303
OY	1225	AGGATTTCTACGTATTAATCTTCCACAGATTTTGGGAATCACATGATGTTTGGCTTGGATC	1284
Db	1304	AGGATTTCTACGTATTAATCTTCCACAGGTTTGGGAATCACATGATGCTCTGGCATTCGATC	1363
OY	1285	ATGTGAAAAGGCATCAAGAATATTAAGTCTGTGTGTATTTAGAAATGGTTTGGAAAT	1344
Db	1364	ATGTGAAAAGGCATCAAGAATATTAAGTCTGTGTGTATTTAGAAATGGTTTGGAAAT	1423
OY	1345	AA 1346	
Db	1424	AA 1425	
RESULT 5			
LOCUS	ABO15426	2139 bp mRNA linear ROD 23-JUN-1999	
DEFINITION	Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, complete cds.		
ACCESSION	ABO15426		
VERSION	ABO15426.1	GI:3702718	
KEYWORDS	Fut9; alpha1,3-fucosyltransferase IX.		
SOURCE	Mus musculus cDNA to mRNA.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus.		
REFERENCE	Kudo,T., Ikehana,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K.		
AUTHORS	1 (sites)		
TITLE	Expression cloning and characterization of a novel murine alpha1,3-fucosyltransferase, mFuc-TIX, that synthesizes the Lewis x (CD15) epitope in brain and kidney		
JOURNAL MEDLINE	9, Biol. Chem. 273 (41), 26729-26738 (1998)		
REFERENCE	98434368		
AUTHORS	2 (bases 1 to 2139)		
TITLE	Kudo,T. and Narimatsu,H.		
JOURNAL	Direct Submission		
FEATURES	Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science, Soka University, Division of Cell Biology: 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan (E-mail:tkudott.soka.ac.jp, Tel:+81-426-91-2495, Fax:+81-426-91-9315)		
Location/Qualifiers			

```

source      1. 2139
            /organism="Mus musculus"
            /db_xref="taxon:10090"
gene         115..1194
            /gene="Fut9"
            115..1194
            /gene="Fut9"
            /codon_start=1
            /product="alpha1,3-fucosyltransferase IX"
            /protein_id="BA33522.1"
            /db_xref="GI:3702719"
            /translation="MTSTSKGLRPLVLCITIGCFMCLLIYIKPTSWPSPESA
            SSVLMKMFSTKIDYFNEETITLVWVPFGQFEDLTSQAFNIGCHLITDRSLYRK
            SHAVLIHRDISMDLTNLPQOARPEPQKIMWNMSPIHTPQKSEIHLFNLTYR
            DSDIQVPEGLFVSTNPEVPSPEKLVWVMSNMEHNAVKYNEIHSISEIHTY
            GOARGEYVNDKNIPLTISTCKEYLSFNSIHDIYETKLYNAFLAGSPVYVGPREN
            YENIPADSFHVEDNPSRELAKYLFKVDNNKNIYLSYPMNRKPTVNLPRFMSHA
            CLACDHVRHREYKSVGNLEKMFN"
polyA_site  2139
            /note="31 a nucleotides"
BASE COUNT  684 a 395 c 368 g 692 t
ORIGIN
Query Match 73.3%; Score 986.8; DB 10; Length 2139;
Best Local Similarity 90.0%; Pred. No. 1,5e-239;
Matches 1069; Conservative 0; Mismatches 117; Indels 2; Gaps 1;
159 CCTAGACACCGATTAGATTAATTAACCAAGATTTGATTAACAGTAGTAGTATA 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
9  CGCTGGATCGCTTGAATGTGATTAACCAAGATTTGATTAACAGTAGTAGTATA 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 ACACTGTCTACGCTGCCATGATGATGCTCTATATATGAAAAATTAAGACATCA 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 GCACATCATCTC--TGCTTCACATGCTATGCTCTACACTGAAAAATTAAGACATCA 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
279 TCCAAAGAAATTTCTGCCCATTTTAAATGTCTGATATATCTGGGCTGTTTACAGCA 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 TCCAAAGCAATTCCTGCCCATTTTAAATGCTGATCATCTCGGCTGCTTCAAGGCA 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 TCTCTCTCATTTACATCAACCTACCAACAGCTGATCTTCAGTCCATGATGATGATC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 TGTCTCTCATTTATATCAAGCCACCAACAGCTGATGATGATGATGATGATGATGAT 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
399 ACCTCTGTGCTGAAAAATGAAAAATCTTTTCCACCACCAAGATTTTAAATGAAC 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 ACCTCTGTGCTGAAAAATGAAAAATCTTTTCCACCACCAAGATTTTAAATGAAC 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 ACTATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 ACCATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 ATGTTCAACATCCAGAGATGCTATCCACACGAGACCTGTTCACTGATGATGATGATGAT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 AGTTCAATATCCAGAGATGCTATCCACACGAGACCTGTTCACTGATGATGATGATGAT 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 GCAATTTCTGATCCATCCAGAGATGCTATCCACACGAGATGATGATGATGATGATGAT 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
427 GCGATTTCTGATCCATCCAGAGATGCTATCCACACGAGATGATGATGATGATGATGAT 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 AGCCACCTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 AGCCACCTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
699 AAGAGTGCATTTGAGCATTTGTTAACTGATCCGATGATGATGATGATGATGATGATGAT 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
547 AAGAGTGCATTTGAGCATTTGTTAACTGATCCGATGATGATGATGATGATGATGATGAT 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
759 CAAGTGCATTTGAGCATTTGTTAACTGATCCGATGATGATGATGATGATGATGATGAT 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
607 CAAGTGCATTTGAGCATTTGTTAACTGATCCGATGATGATGATGATGATGATGATGAT 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
819 AAGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 667 AAGAGAACTTGGTGTGCTGGGTTGTGATTAACCTGGAACCTGAGATCCAGGTCAG 726
Qy 879 TATTACATAGACTAAGCAAAAGCATTTGAATTCATACCTAGCGGCAACATTTGAGAA 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 727 TATTACATAGACTAAGCAAAAGCATTTGAATTCATACCTAGCGGCAACATTTGAGAA 786
Qy 939 TATGTCATGATTAATAATTTGATTCCTACCATATCTGCTTTGTAATTTATCTTCCTTT 998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 787 TACGTGAACGATTAATAATTTGATTCCTACCATATCTGCTTTGTAATTTATCTTCATTT 846
Qy 999 GAAATTCATTCACAGAGATTAACATACAGGAAACCTTACAAATCTTTCTGCTGCC 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 847 GAAATTCATTCACAGAGATTAACATACAGGAAACCTTACAAATCTTTCTGCTGCC 906
Qy 1059 TCTGTACTGTTGTTCTGCGACCATCTAGGAAACATATGAGATTAATTCACAGAGAT 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 907 TCAGTATCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Qy 1119 TCATTCATTCATGTTGAGATTAATTAACCTCCAGTACGAGTACGAAAGTATGAGAA 1178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 967 TCATTCATTCATGTTGAGATTAATTAACCTCCAGTACGAGTACGAAAGTATGAGAA 1026
Qy 1179 GTTCGAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1027 GTTCGAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1086
Qy 1239 AATCTTCACAGATTTGGAATCAATCAATGATGTTGCTTGCATCATGTAAGAAAGCAT 1298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1087 AACCTTCACAGATTTGGAATCAATCAATGATGTTGCTTGCATCATGTAAGAAAGCAT 1146
Qy 1299 CAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1147 CAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
AF345993 1128 bp mRNA linear ROD 07-MAR-2001
LOCUS AF345993
DEFINITION Rattus norvegicus alpha1,3-fucosyltransferase IX (Fut9) mRNA,
complete cds.
ACCESSION AF345993
VERSION AF345993.1 GI:13242183
KEYWORDS SOURCE
ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1128)
Baboval,T., Henion,T., Kinnally,E. and Smith,F.I.
Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX)
and comparison of the expression of fuc-TIV and fuc-TIX genes
during rat postnatal cerebellum development
J. Neurosci. Res. 62 (2), 206-215 (2000)
11020213
JOURNAL PUBMED 2 (bases 1 to 1128)
AUTHORS Smith,F.I. and Baboval,T.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver
Center, 200 Trapelo Rd., Waltham, MA, USA
LOCATION/Qualifiers
FEATURES
source 1..1128
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1128
/gene="Fut9"
/gene="Fut9"
49..1128
/gene="Fut9"
/codon_start=1
/product="alpha1,3-fucosyltransferase IX"
/protein_id="AK1591.1"
/db_xref="GI:13242184"

```


OY	500	CCTACATCCTGCGCAAGCATATTTCACAATCCAAAGATGCCATCTCCAMAAGCACGGTTC	559
Db	501	TCTAACACTCTTGCCCAACGAATGTTCACACATCCATGGGTGCATCTGTGATTTGACCGCTC	560
OY	560	ACGTGTACACAATAATCCCATGCAAGTCTGTGATTCATCACCGAGACATCAGTTGGATTCGAC	619
Db	561	ACTAATATTAACAAGTCCCCATCAGTCTGTTCATACATCAACAGGAGCAATTAACGGGATCTGAC	620
OY	620	AAATTTACGTCAGCAAGCTTAGGCCACCCTTCCAGAAATGGATTTGGATTTGGATTCGAAATC	679
Db	621	TAATTTACCTCAGCACMACCCAGGCCACCATTCCAGAAATGGATTTGGATTTGGATTCGAAATC	680
OY	680	ACCAACTGACACACTCCCCMAAAGAAGTGGCATTTGAGCACTTTGTAACTCGACTGTGACTTA	739
Db	681	TCCAATCTCAATACACCAACAAAAGAGTGGCATTTGAAACACCTTTTAACCTAACCTCGACTTA	740
OY	740	CCGCCCTGATTCAGAAATCCAAATGGCTTTATGGCTCTTATGACGCTTAAGCAACAAATCCCTT	799
Db	741	CCGGCTGATTCAGATATTTCAGTGGCTTATGGCTTCATATGATTTGGTTCACAGCGCCTT	800
OY	800	CGGTTTGAAGTCCCAAGCAAGAAGAAATTTGGTGTCTGGTGTGTAGTAACTCGAACC	859
Db	801	TACATATGAAGTCCCAAGTAAAGAAATTTGGTTTTGTTGGTTGTAAATGTAACGGAACCC	860
OY	860	TGAGCATGCGCAAGTCAAGTATTTACAAATGAGCTTAACCAAAAGCATGAAATCCATACCTTA	919
Db	861	TGAGCATGCTCGAGTCAAGTAAATTTACAAATGAGCTTAACCAAGTCAATGAAATCCATACCTTA	920
OY	920	CGGGCAGACATTTGGAGAAATGTGTCAATGATAAAAAATTTGATTCCTACATATCTGCTTG	979
Db	921	CGGGCAGACCTTTGGAGACTAGCTCAACGATTA AAACTTAATTCACAACATCTCCAGTTG	980
OY	980	TAAATTTATCTTCCPTTGGAAAATTCAAATCCACAAGGATTTACATACCGGAAAAAGCTATA	1039
Db	981	CAAATTTCACTTCAATTTGAAAATTCAAATCCACAAGGATTTACATTAACGAGAAACCTTTA	1040
OY	1040	CAATGCTTTTGGCGCTGTGTACTCTGTGTCTGTGTGGAGACATCTAGGAGAAAACATATGA	1099
Db	1041	CAATGCTTTTGGCGCGATCATGATACCAATTTGTACTGGGCCCTCCAGAGAAAACATATGA	1100
OY	1100	GAAATTAATTCAGCAGATTCATCTCTTATGTGGAAATTTAACTCTCCAGTBAAGCT	1159
Db	1101	GAATTAATTCAGCAGGAGACTCTTTATTCACAGTGGAAATTTCTATCCCCAGAGAGCT	1160
OY	1160	AGCAAGTATTCGAAGGAGTGCACAAAACAATAAGTAAATACCTTAAGTACTTTAACTG	1219
Db	1161	TCCAGAAATCTCTTGATGCTGTGACAAAAATAACAAGATGTACCTTAGTTACTTCAACTG	1220
OY	1220	GAGGAAAGATTCACATCGTAAATCTTCCAGATTTTGGGAATACATGCATGTTGGCTTG	1279
Db	1221	GAGGAAAGATTTCTCAGTGCATCTTCTTAATATCTGGGAATCTCAATGCATGTCTGCTTG	1280
OY	1280	CGATCATGTGAAAAGGCATCAAGAAATATATAGTCTGTGGTAAATTTAGAGAAATGGTTTTG	1339
Db	1281	TGATCATGTGAAAAGACACCGAAGATACAAATGTGAATTTAGAAAATTTAGAAATTTGGTTTTG	1340
OY	1340	GAATTTAA 1346	
Db	1341	GAACCTAA 1347	

RESULT 9					
AB035905	AB035905	1080 bp	DNA	linear	VRT 05-JAN-2002
LOCUS	Xenopus laevis gene for xrt9, complete cds.				
DEFINITION	AB035905				
ACCESSION	AB035905.1	GI:18146863			
VERSION	xrt9...				
KEYWORDS	Xenopus laevis DNA.				
SOURCE	Xenopus laevis				
ORGANISM	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;				

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL	COMMENT FEATURES	source
Xenopodinae: <i>Xenopus</i> . 1 (sites) Kaneko, M., Nishihara, S., Kitano, T., Narimatsu, H. and Saitou, N. The evolutionary history of glycosyltransferase genes unpublished 2 (bases 1 to 1080) Kaneko M., Saitou, N. and Kitano, T. Direct Submission		
Submitted (17-DEC-1999) Mika Kaneko, National Institute of Genetics, Laboratory of Evolutionary Genetics, Yata 111, Mishima, Shizuoka 411-8540, Japan (E-mail:mkaneko@med.id.yamagata-u.ac.jp, Tel.:81-539-81-6790, Fax:81-559-81-6789) Sequence updated (29-Feb-2000). Location/Qualifiers I. 1080		

```

gene
/gene-"xFuc-Tix"
1..1080
CDS
1..1080
/gene-"xFuc-Tix"
/codon_start-1
/product-"xftug"
/protein_id-"BAB82486.1"
/db_xref-"GI:16146864"
/translation-"MTSASGILRAFLAVSGVLCIMICILLIYKPKPTNMNISSPIESS
HSAVLKRNPFSSKONTYNEVQQLIMWVPEFGOTPELKSCOTLFNYGCHLFTDALAN
SHAVLLHHRDDISMDLILNLPQQLRPPQOKWIMMNLSEPHHPQKSGIHLFNLTLYR
DSDIQVYFVWVSSTRPPDEEYVPSKDKIVCWVNMNPNENHARVRYVYMLKLTITIT
GAGFYSLRDKSLPLPTIASCKFYSLFENSIKHDKITKRLNALLAGSVPLVLPDPREE
YENYVLPADSFIVHEDFLSPRELADHLLNLNMDTQYLSYNNRKHFTVNMPIHWMESH
CLACBCHKHQKYSVGNLMEKFWNA"
332 a 231 c 202 g 315 t
BASE COUNT
ORIGIN

```

Query Match	Similarity	Score	DP	Length
Best Local	76.7%	676	8	1080
Matches	828	Conservative	0	Mismatches 255; Indels 0; Gaps 0

Query	Match	Similarity	Score	DP	Length
267	ATGCAATCAATCAACAAAGAAATCTCTGCCCATTTTAAATGTCTGCATTATCCGAGC	326			
Db	1 ATGCAATCAATCAATCAACAAAGAAATCTCTGCCCATTTTAAATGTCTGCATTATCCGAGC	60			
QY	327 TGTTCATGCGATGCTCTCTCATTTACATCAAACTTACCAACGCTGATCTTCAGTCA	386			
Db	61 TGTTCATGATGATGTTTACTGATATACGTCAAACCAACAACTGATTTCCAGTCT	120			
QY	387 ATGGAATCAGCCAGCTGCTGTGCTGAAATGAAAACTTCTTTCACCAAAATGATAT	446			
Db	121 ATGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	180			
QY	447 TTTAATGAATCTACTATTCGTGGTGGGTGTGGCCATTTTGGGAGACCTTGGACTTACA	506			
Db	181 TACAAATGAACCCGCTCTTCTTAATTTGGGGTGGCCCTTTTGGGAGACTTTTGAACCTAAA	240			
QY	507 TCTGCGCAAGCATGTTCAACATCCAAAGATGTCATTCACAAGGACCGTTCACTGAC	566			
Db	241 TCTGCGCAAGCATGTTTAAATTTATGGTGGCCATTTACTACGACCGGCACTGTAC	300			
QY	567 AACAAATCCATGCACTTGTGATTCATCAGCAAGCATCACTGGGATGTGCAAAATTA	626			
Db	301 AACAAATCCATGCTCTCTCTTCTTACACCAAGCAATTTACTGTGGATGTCAATTA	360			
QY	627 CCTCAGCAAGCTGAGCGCCACTTCCAGAAATGATTTGGATGAAATTTGAAATCCCACT	686			
Db	361 CCCCACAGCTGAGCGCCACTTCCAGAAATGATTTGGATGAAATTTGAAATCCCACT	420			
QY	687 CAACATCCCCAAAAGATGGCATTTAGACACTTGTAACTGTACTGTACTTACGCCGT	746			
Db	421 CATACGCAACAAAAGATGGCATTTAGACACTTGTAACTGTACTGTACTTACGCCGT	480			
QY	747 GATTGATATCAAGATGCTTATGGCTTGTGACGATACCAAAATCCCTTCGTGTT	806			

Db 481 GACTCAGATATCCAGTAGTGCCTTATGCTTCATGCTGTAAGCAGACAAAACATTGACTTT 540
QY 807 GAAGTCCCAACCAAGAGAAATGGTGTGCTGGTTGTGAGTAACTGAGACCCAGACAT 866
Db 541 GAGGTGCCCAACCAAGAGAAATGGTGTGCTGGTTGTGAGTAACTGAGACCCAGACAT 600
QY 867 GCCAGAGCAGATATATACATGAGTAAAGCAAAAGCAATGAAATTCATACCTAGGGCAA 926
Db 601 GCCAGAGCAGATATATACATGAGTAAAGCAAAAGCAATGAAATTCATACCTAGGGCAA 660
QY 927 GCATTTGAGATATGTCATGATATAAATTTGATTCCTACCATATGCTGTAATTT 986
Db 661 GCTTTGGTGAATACCTGAGGAGATATAAGCTGCTTCCACTATTCATCTTGCAAAATTT 720
QY 987 TATCTTCTCTTGAATATTCATATCCAGAGATTCATACATGAGGAAAGCTATACATGCT 1046
Db 721 TATCTGCTCTTGAATATTCATATCCAGAGATTCATACATGAGGAAAGCTATACATGCA 780
QY 1047 TTTTGGCTGGCTGTGACCTGTTGTTCTTGAGGACCATTCAGGAAACTATGAGATTTAT 1106
Db 781 TTAATGGCTGGCTGTGACCTGTTGTTCTTGAGGACCATTCAGGAAACTATGAGATTTAT 840
QY 1107 ATTCCAGAGATTCATATTCATATCCAGAGATTCATACATGAGGAAAGCTATACATGCT 1166
Db 841 ATTCTGCTGAGATTCCTTATATCCAGTGAAGATTTTCTCTCCAGAGATTTAGCTGAC 900
QY 1167 TATCTGAGAGAGTGCAGAAAAACAAATAGTTATACCTTACTTACTGAGAGAG 1226
Db 901 CATCTTTAATGCTGAGAACAAAGATACAGAACAGTACCTGAGCTACATGAGAGGAAA 960
QY 1227 GATTCACATGTAATTCCTCCAGATTTTGGGATCAGATGATGCTTGGCTGGATCAT 1286
Db 961 CATTTACAGTCAATATGCTCCTGCTGAGGAAATCCATGATGCTTACCTGAGCAT 1020
QY 1287 GTGAAAAGGCATCAGAGATATAGTGTGCTGTAATTTAGAGAAATGTTTGAATTTAA 1346
Db 1021 GTGAAAAGGCATCAGAGATATAGTGTGCTGTAATTTAGAGAAATGTTTGAATTTAA 1080

RESULT 10
AC100674 71396 bp DNA linear HTG 22-NOV-2001
LOCUS
DEFINITION Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.
AC100674
VERSION AC100674.1 GI:17048040
KEYWORDS HTG; HTGS_PHASED.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 71396)
Biren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-167K24
Unpublished
2 (bases 1 to 71396)
Biren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barina,N., Bastien,Y., Boguslavsky,L., Bouknight,B.,
Brown,A., Camarata,J., Campoliano,A., Chang,Y., Chazaro,B.,
Cooke,P., DeArliano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagde,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labrecque,K.,
Lamasares,R., Landers,T., Lehoczeky,J., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Melidiri,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olivier,J., Peterson,K., Phunhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,J., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

TITLE
JOURNAL
COMMENT

Strauss,N., Subramanian,A., Talmaz,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15932
Center clone name: 167_K_24

* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved. 788: contig of 788 bp in length
* 1
* 789 888: gap of 100 bp
* 889 1702: contig of 814 bp in length
* 1703 1802: gap of 100 bp
* 1803 2642: contig of 840 bp in length
* 2643 2742: gap of 100 bp
* 2743 3637: contig of 895 bp in length
* 3638 3737: gap of 100 bp
* 3738 4635: contig of 898 bp in length
* 4636 4735: gap of 100 bp
* 4736 5657: contig of 922 bp in length
* 5658 5757: gap of 100 bp
* 5758 6692: contig of 935 bp in length
* 6693 6792: gap of 100 bp
* 6793 7703: contig of 911 bp in length
* 7704 7803: gap of 100 bp
* 7804 8742: contig of 939 bp in length
* 8743 8842: gap of 100 bp
* 8843 9787: contig of 945 bp in length
* 9788 9887: gap of 100 bp
* 9888 10819: contig of 932 bp in length
* 10820 10919: gap of 100 bp
* 10920 11808: contig of 889 bp in length
* 11809 11908: gap of 100 bp
* 11909 12830: contig of 922 bp in length
* 12831 12930: gap of 100 bp
* 12931 13847: contig of 917 bp in length
* 13848 13947: gap of 100 bp
* 13948 14886: contig of 939 bp in length
* 14887 14986: gap of 100 bp
* 14987 15911: contig of 925 bp in length
* 15912 16011: gap of 100 bp
* 16012 16903: contig of 892 bp in length
* 16904 17003: gap of 100 bp
* 17004 17906: contig of 903 bp in length
* 17907 18006: gap of 100 bp
* 18007 18931: contig of 925 bp in length
* 18932 19031: gap of 100 bp
* 19032 19945: contig of 914 bp in length
* 19946 20045: gap of 100 bp
* 20046 20951: contig of 906 bp in length
* 20952 21051: gap of 100 bp
* 21052 21949: contig of 898 bp in length

21950 22049: gap of 100 bp
22050 22950: contig of 901 bp in length
22951 23050: gap of 100 bp
23051 23945: contig of 895 bp in length
23946 24045: gap of 100 bp
24046 24955: contig of 910 bp in length
24956 25055: gap of 100 bp
25056 25980: contig of 925 bp in length
25981 26080: gap of 100 bp
26081 27052: contig of 972 bp in length
27053 27152: gap of 100 bp
27153 28067: contig of 915 bp in length
28068 28167: gap of 100 bp
28168 29076: contig of 909 bp in length
29077 29176: gap of 100 bp
29177 30093: contig of 917 bp in length
30094 30193: gap of 100 bp
30194 31099: contig of 906 bp in length
31100 31199: gap of 100 bp
31200 32097: contig of 898 bp in length
32098 32197: gap of 100 bp
32198 33113: contig of 916 bp in length
33114 33213: gap of 100 bp
33214 34125: contig of 912 bp in length
34126 34225: gap of 100 bp
34226 35133: contig of 908 bp in length
35134 35233: gap of 100 bp
35234 36150: contig of 917 bp in length
36151 36250: gap of 100 bp
36251 37171: contig of 921 bp in length
37172 37271: gap of 100 bp
37272 38159: contig of 888 bp in length
38160 38259: gap of 100 bp
38260 39156: contig of 897 bp in length
39157 39256: gap of 100 bp
39257 40188: contig of 932 bp in length
40189 40288: gap of 100 bp
40289 41210: contig of 922 bp in length
41211 41310: gap of 100 bp
41311 42222: contig of 912 bp in length
42223 42322: gap of 100 bp
42323 43218: contig of 896 bp in length
43219 43318: gap of 100 bp
43319 44233: contig of 915 bp in length
44234 44333: gap of 100 bp
44334 45246: contig of 913 bp in length
45247 45346: gap of 100 bp
45347 46261: contig of 915 bp in length
46262 46361: gap of 100 bp
46362 47279: contig of 918 bp in length
47280 47379: gap of 100 bp
47380 48276: contig of 897 bp in length
48277 48376: gap of 100 bp
48377 49282: contig of 906 bp in length
49283 49382: gap of 100 bp
49383 50305: contig of 923 bp in length
50306 50405: gap of 100 bp
50406 51314: contig of 909 bp in length
51315 51414: gap of 100 bp
51415 52335: contig of 921 bp in length
52336 52435: gap of 100 bp
52436 53163: contig of 728 bp in length
53164 53263: gap of 100 bp
53264 54210: contig of 947 bp in length
54211 54310: gap of 100 bp
54311 55199: contig of 889 bp in length
55200 55299: gap of 100 bp
55300 56212: contig of 913 bp in length
56213 56312: gap of 100 bp
56313 57287: contig of 975 bp in length
57288 57387: gap of 100 bp
57388 58245: contig of 858 bp in length
58246 58345: gap of 100 bp

58346 59236: contig of 891 bp in length
59237 59336: gap of 100 bp
59337 60272: contig of 936 bp in length
60273 60372: gap of 100 bp
60373 61289: contig of 917 bp in length
61290 61389: gap of 100 bp
61390 62299: contig of 910 bp in length
62300 62399: gap of 100 bp
62400 63322: contig of 923 bp in length
63323 63422: gap of 100 bp
63423 64336: contig of 914 bp in length
64337 64436: gap of 100 bp
64437 65354: contig of 918 bp in length
65355 65454: gap of 100 bp
65455 66352: contig of 898 bp in length
66353 66452: gap of 100 bp
66453 67362: contig of 910 bp in length
67363 67462: gap of 100 bp
67463 68370: contig of 908 bp in length
68371 68470: gap of 100 bp
68471 69379: contig of 909 bp in length
69380 69479: gap of 100 bp

Query Match 47.5%; Score 639; DB 2; Length 71396;
Best Local Similarity 85.0%; Pred. No. 8e-151;
Matches 764; Conservative 0; Mismatches 122; Indels 13; Gaps 4;

OY 338 ATGCTCTCATTTACATCAACCTACCAACGCTGATCTTCAGTCCATGGAATCAGC 397
DB 65456 ATCTCTGCCCTGTGTTGATTTCCACCAACGCTGCTTCTAGTCCATGAGTGTGC 65515
OY 398 CAGCTCTGCTGAAAGAAAGAACTCTTTCCACCAACGATATTTATGAAAC 457
DB 65516 AAGTCTGTGCTGAAAGAAAGAAATTTCTTCTCCAAAGATATTTATGCAAGC 65575
OY 458 TACTATCTGTGTGGGTGGGTGGCATTTGGGACACCTTGACCTTCATCTGCCAAGC 517
DB 65576 TACATTTCTGTGTGGGTGGGTGGCATTTGGGACACCTTGACCTTCATCTGCCAAGC 65635
OY 518 AATGTCAACATCCAAAGATGCCATCTCACACGAGACCTTCACTACCAAAATCCCA 577
DB 65636 AATGTCAACATCCAAAGATGCCATCTCACACGAGACCTTCACTACCAAAATCCCA 65695
OY 578 TGCACTTGTATCCATCACCGACATCACTGTGGATCTGACAATTTACCTGAGAAGC 637
DB 65696 TCGGCTCTGATTCACATCAGATGACATCAGCTGGATCTGACTACCTACCTGAGCAGGC 65755
OY 638 TAGGCACCCCTTCCAGAAATGGATTTGGATGAATTTGGAATCCAACTCAGCTCCCA 697
DB 65756 CAGGCCACCCCTTCCAGAAATGGATTTGGATGAATTTGAGATCCACCACCTCAGCTCCCA 65815
OY 698 AAAGAGTGGCATTTAGCACTTGTTAACCTGACCTGACTTACCGGCGTATCAATATN 757
DB 65816 AAAGAGTGGCATTTAGCACTTGTTAACCTGACCTGACTTACCGGCGTATCAATATN 65875
OY 758 CCAAGTGGTATGAGCTTGTGAGTGGATGAGTGGATGAGTGGATGAGTGGATGAGTGG 817
DB 65876 CCAAGTGGTATGAGCTTGTGAGTGGATGAGTGGATGAGTGGATGAGTGGATGAGTGG 65935
OY 818 CAAAGAGAAATTTGCTGTGCTGGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 877
DB 65936 CAAAGAGAAATTTGCTGTGCTGGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 65995
OY 878 GTATTACAAATGATTAAGCAAAAGATGGAATTCATATCCATGAGGCAAGCATTTGAGA 937
DB 65996 GTATTACAAAGATGATTAAGCAAAAGATGGAATTCATATCCATGAGGCAAGCATTTGAGA 66055
OY 938 ATATGCAATGATTAATTTGATTTCTACCATATCTGCTGTAATTTATTTATTTCTT 997
DB 66056 ATATGCAATGATTAATTTGATTTCTACCATATCTGCTGTAATTTATTTATTTCTTCT 66115
OY 998 TGAATTTCAATCCACAAGGTTTACATCAGGAAAGCTTACAAATGCTTTCTGTGCTGG 1057
DB 66115 TGAATTTCAATCCACAAGGTTTACATCAGGAAAGCTTACAAATGCTTTCTGTGCTGG 1057

D6 66116 GAAACATCATTCACAATTAATCTCCCAAG-----TCACATGCTTTTGGCTGG 66166

OY 1058 CTGTACCTGTTGTTGGACCATCTAGGAAACATAGATAATATATTCACAGAGA 1117
|||||
D6 66167 TTCACTACCTGTTGTTGGACCATCTAGGAAACATAGATAATATATTCACAGAGA 66226
|||||

OY 1118 TTCACTACCTGTTGTTGGACCATCTAGGAAACATAGATAATATATTCACAGAGA 11174
|||||

D6 66227 TTCACTACCTGTTGTTGGACCATCTAGGAAACATAGATAATATATTCACAGAGA 66286
|||||

OY 1175 GGAAGTCGACAAAACATAGATT-ATACCTTACTTACTTACTGAGAGAGATTTC 1232
|||||

D6 66287 GGAAGTCGACAAAACATAGATTGTTGTTACTTACTTACTGAGAGAGATTTC 66345
|||||

RESULT 11
AB023627 1567 bp DNA linear VRT 02-APR-1999
LOCUS
DEFINITION
Dantio rerio zrf11 gene for alpha(1,3)fucosyltransferase, complete
cds.
AB023627.1 GI:4587295
VERSION
alpha(1,3)fucosyltransferase.
KEYWORDS
Dantio rerio DNA.
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Dantio.
1 (sites)

REFERENCE
AUTHORS
TITLE
JOURNAL
Kageyama, N., Natsuka, S. and Hase, S.
Molecular cloning and characterization of two zebrafish
alpha(1,3)fucosyltransferase genes developmentally regulated in
embryogenesis
J. Biochem. (1999) In press
2 (bases 1 to 1567)
Kageyama, N., Natsuka, S. and Hase, S.
Direct Submission
Submitted (12-FEB-1999) Shunji Natsuka, Osaka University, Graduate
School of Science; Machikaneyama 1-1, Toyonaka, Osaka 560-0034,
Japan (E-mail:natsuka@chem.sci.osaka-u.ac.jp, Tel:81-6-6850-5381,
Fax:81-6-6850-5383)

FEATURES
source
location/Qualifiers
1..1567
/organism="Dantio rerio"
/db_xref="taxon:7955"
177..1352
/gene="zrf11"
177..1352
/gene="zrf11"
/codon_start=1
/product="alpha(1,3)fucosyltransferase"
/protein_id="BA076706.1"
/db_xref="GI:4587295"
/translation="MDKILTPSKAAQVITLPMILSTVCFYVYNNPTTFKFP
ADVHTNCSCTETCLDVLYKONKCTIKNSDPSSTTPANRPAKAKDDQETILLI
WVYFGARPDVGVGSGQFNGCHLDHSMIGKAGVMEHNDLSADLPQPPRPAFO
KIMWMMSPSNHNSPNSLINDGNLSTSPKSDIDIVPYGRLIDADEQKNFTIPKRD
KLVCWIVSNQEHKRSQYNELVKIKIKEYAGGHNPNRVNDQDGVSSCKEYLSF
ENSHRDYFTEKLFNPLALGTVPVYLPSPSRDNEEFLPRDAFIHVDDEPRKELADHL
KSLDNEEDYRQFPMNRKHFISMSSSFLGHACRTCDYLKRNKHYAIDKLGKFWNG"

BASE COUNT 499 a 334 c 293 g 441 t
ORIGIN

Query Match 20.6%; Score 277.8; DB 5; Length 1567;
Best Local Similarity 58.5%; Pred. No. 1.3e-59;
Matches 527; Conservative 0; Mismatches 362; Indels 12; Gaps 2;

OY 449 TAATGAACACTATCTGCTGCTGGTGGCCATTTGGCGACCTTTGACCTTACATC 508
|||||
D6 461 TCAAGAAACCAATCTGTTGATCTGGGTATACCTTGGAGACAGCTTTGACTGGGAGT 520
|||||

OY 509 CTGCCAAGCATGTTCAACATCCAGAGATGCTCACAACGACCGCTTACCTGTTCAA 568
|||||

D6 521 GTGTGTTGCGAGTTCAACATCCAGCGCTGTCATTTAACAGACAGCAAGATATGATCCA 580

OY 569 CAATTCATGCGAGTTGCTGATCATCACAGACATCATGTTGGATGTCAGAAATTTACC 628
|||||

D6 581 GAAAGCGCATGAGATATGTTTACACAGAGATCTAGGCGCG-----ATTGGCC 631
|||||

OY 629 TCAGACAGCTAGGCGACCCCTCCAGAAATGATTTGGATGATTTGGATTCACCACTCA 688
|||||

D6 632 GCAACCCCTTCGACACACTTCCAAAAGTATGTTGTAATGATGAGTGTCCGAGTAA 691
|||||

OY 689 CACTCCCAAAAGAGTGGCATTTGACACTTGTTAACCTGATCTGATTAACCCCTGTA 748
|||||

D6 692 TTCCATCCCAACAGCTTGATTAATGATGTTTAACTCCGTCAGTTCGCAAGGA 751
|||||

OY 749 TTCAATATCCAAAGTGCCTTA---TGCTCTTGACGGTATGACAAATCCCTGCTGT 805
|||||

D6 752 TTCAATATCCAAAGTGCCTTAAGCTTACGCTGATGATGATGATGATGATGATGATGAT 811
|||||

OY 806 TGAAGTCCCAAGCAAGAGAAATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 865
|||||

D6 812 CACCATCCCAAAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 871
|||||

OY 866 TGCCAGAGTCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925
|||||

D6 872 CAACAGTTCACAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 931
|||||

OY 926 AGCATTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
|||||

D6 932 GCACCTTCAATACCGGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 991
|||||

OY 986 TTATCTTCTCTTTGAAATTTCAATCAACAGAGATTTACATCAGGAGAAAGCTATACAAATGC 1045
|||||

D6 992 CTACCTTCTCTTTGAAATTTCAATCAACAGAGATTTACATCAGGAGAAAGCTATACAAATGC 1051
|||||

OY 1046 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1105
|||||

D6 1052 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1111
|||||

OY 1106 TATTCACAGAGATTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1165
|||||

D6 1112 TATTCACAG 1171
|||||

OY 1166 GTATCTGAAGAGATGCAACAAAACAAATTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA 1225
|||||

D6 1172 TCACCTTAAATCTTGGACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1231
|||||

OY 1226 GGAATTCACGTAATCTTCCACAGATTTGGGATTCACATGATGATGATGATGATGATGAT 1285
|||||

D6 1232 ACATTTTATTTCAATGAGTTCATCTATTTGGGCTTGAACATGCTGTAGGACTGTGTGATTA 1291
|||||

OY 1286 TGTGAAAAGCATCAAAATTAAGTCTGTTGTAATTTAGAGAAATGATTTGGATTTA 1345
|||||

D6 1292 TCTTAAAAGGAATTAAGACATTAATGATTTAAAGATCTGAAAGCTGTGTTTGGGTTA 1351
|||||

OY 1346 A 1346
|||||

D6 1352 A 1352
|||||

RESULT 12
AB023628 1832 bp DNA linear VRT 02-APR-1999
LOCUS
DEFINITION
Dantio rerio zrf2 gene for alpha(1,3)fucosyltransferase, complete
cds.
AB023628
VERSION
AB023628.1 GI:4587297
KEYWORDS
alpha(1,3)fucosyltransferase.
SOURCE
Dantio rerio DNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Dantio.
1 (sites)

FEATURES	source
gene	gene
CDS	CDS
ORIGIN	581 a 712 c 818 g 546 t
BASE COUNT	
ORIGIN	581 a 712 c 818 g 546 t

Query Match	13.1%;	Score 176.2;	DB 5;	Length 2657;
Best Local Similarity	-53.1%;	Pred. No. 9.2e-34;		
Matches 451; Conservative	0;	Mismatches 383;	Indels 15;	Gaps 3

OY	459	ACATTCCTGGTGGTGGTGGTGGCCATTGGGGACACCTTTACCTTACATCTGACCCACAGA	518
OY	459	ACATTCCTGGTGGTGGTGGTGGCCATTGGGGACACCTTTACCTTACATCTGACCCACAGA	518
Db	341	ACGGGCTGCTGTGTGGGAGACCTTTGGCGCGCCCTTGGGGCCCGGCGGACGTGCGGAG	400
OY	519	ATGTTCAACATCCAAAGATGCGATCTCAACAAGGACCGCTTCACTGTACACAATCCAT	578
Db	401	CGCTACCAATACACGGGGTGGCTCTCTAGCGCCGACCGCGGGGGGTAGCGGAGGTCTGG	460
OY	579	GCAGTTTGCATCCATCCACCGAGACATCAGTTGGGATCTGCAAAATTTACCTAGCAAG--	636
Db	461	GCCGTGCTCTTCCACACCGGGACCTTGGCCTGCACGGCCGCCAGGGGGCTGCCGCCGGG	520
OY	637	-----CTAGGCCACCCCTTCCAGAAATGGATTGGATGAATTTGGAAATCCACACTCAC	689
Db	521	CCCCCGCGCACCCCGCGCGCACCGGTGGGTGGATGATGAATCTTCAAGTCCGCTCGCAC	580
OY	690	ACTCCCAAAAGATGCGCATTTGAGACATTTTAACTGACTCTGACTTACCGCGCTGAT	749
Db	581	TCCCGCGGGGTGCGGGGCTGGCCGGCTCTTCAACTGAGCAATGTCGTACCGACGGAG	640
OY	750	TCAGATATCCAAAGTGCCTTTTATGGCTTTTACGCGTAAAGCAAAATCCCTGCTGTGAA	809
Db	641	TCGGACGCTTCTGTGGCTACGGGTACTCTTACGAGCGCGCGTGGCGCGCCCTTCTGTG	700
OY	810	GTCGCACGCAAAAGAAATGGTGTGGTGGTGTAGTATCTGAAATCCCTGACATGCG	869
Db	701	CTGGCTGCAAAACCGCGGTGGTGGCTGGGTCTATCAGCAATCGAATCGAAGACGACGC	760
OY	870	AGAGTCAGATTTTCATATGAGCTTACGAAAGCATTTGAAATCCATACCTACGAGGCAAGCA	929
Db	761	GCGTGGCTCTACACGGGACACTGTAAAGGACACTGCGCATGACGCTGTACGGG--GGG	817
OY	930	TTTGGAGATATGTCAATGATTAATAAATTTGATTCTCCACATATCTGCTGTAAATTTAT	989
Db	818	CGCGGATGGCGCTGTGGAGGGGAGCGTGTAAAGCGGTGCGGCTCCACAAAGTTCTAC	877
OY	990	CTTTCTTTGAAATTTCAATCAATCCACAAAGATTACATACGAGAAAAGCTAT--ACAATGCT	1046
Db	878	CTGGCCTTCGGAATCTCCAGCACACGAGCATACATACCGAAGAAAGCTGTGGAAAGACGC	937
OY	1047	TTTTGGCTGCTCTGTACCTGTGTGTTCGAGCAATCTAGGGAANAACATTAAGATATT	1106
Db	938	TTTCCCGCGACGCGGGTGGCTGTGTGTCCTGCGGCCCGGACGAGGCCAATACAGACGCTTC	997
OY	1107	ATTCCAGACAGATTCATTCATTCATATGTGGAAGATTATATACCTCCCGAGTAGCTACGAAG	1166
Db	998	ATCCCCCGGACCTCTTCATCCACAGTGAAGATTTCCCGACGCCACGAGCTGCTGGCGACC	1057
OY	1167	TATCTGAAGGAAGTCGACAAAAACAAATAAATATTATACCTAGTTACTTTAACTGAGGAG	1222
Db	1058	TACTGAAATTCCTCGATTAATAAACAAAGCCAGCTACAGGAGATATTTCCCTCGGGGAAAC	1112
OY	1227	GATTCACTGAAATCTTCCACGATTTTGGGAATCAATGCAATGTTGGCTGCGCATCAT	1286
Db	1118	AAGATATGAATCCACGTCACGCTCTTCTTGGGATGACATTTACTCGCAAGTTTTCGAGGCC	1177
OY	1287	GTGAAAGG 1295	
Db	1178	GTGAGAGG 1186	

```

Search completed: October 6, 2002, 04:46:41
Job time : 1826.68 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 151.049 Seconds

(without alignments)
15299.444 Million cell updates/sec

Title: US-09-744-748-5_COPY_29_1374

Perfect score: 1346
Sequence: 1 actgctctccctcgcagctc.....agaaatggtttgtaattaa 1346

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1346	100.0	2676	21	AAZ92647 Human alpha-1,3-fu
2	1088	80.8	2822	21	AAZ92646 Human alpha-1,3-fu
3	986.8	73.3	2170	21	AAZ92645 Murine alpha-1,3-f
4	939.2	69.8	2036	21	AAZ92672 Murine alpha-1,3-f
5	127.4	9.5	1814	18	AAT59505 Murine myeloid-lin
6	127	9.4	3594	18	AAT85219 Mouse alpha-fucosy
7	114.4	8.5	1701	15	AAQ77332 Human alpha-1,3-f
8	114.4	8.5	1701	21	AAF21136 Human low adenosin
9	114.4	8.5	1701	21	AAA35014 Human adenosine re

10	114.4	8.5	6944	21	AAF21137 Human low adenosin
11	114.4	8.5	6944	21	AAA35015 Human adenosine re
12	113.4	8.4	1654	15	AAQ56905 pCDNA1-alpha-(1-3)
13	113.4	8.4	1654	18	AAT61680 Human alpha(1,3)-f
14	112.2	8.3	1086	15	AAQ56912 Lewis blood group
15	112.2	8.3	2042	18	AAT80111 Fucosyltransferase
16	112.2	8.3	2043	12	AAQ13300 Glycosyltransferas
17	112.2	8.3	2043	15	AAQ56906 DNA encoding a gly
18	112.2	8.3	2043	18	AAT76769 Human alpha 1,3/4
19	112.2	8.3	2043	18	AAT61675 Human alpha(1,3)/1
20	111.8	8.3	1126	21	AAF21135 Human low adenosin
21	111.8	8.3	1126	21	AAA35013 Human adenosine re
22	111.8	8.3	1316	15	AAQ56911 DNA encoding a gly
23	111.8	8.3	1316	18	AAT61679 Human alpha(1,3)-f
24	103.6	7.7	795	17	AAT13798 Alpha-(1,3/1/4)-fu
25	102.4	7.6	2175	11	AAQ06691 CDX, a MTLA involv
26	100.8	7.5	1256	21	AAF21134 Human low adenosin
27	100.8	7.5	1256	21	AAA35012 Human adenosine re
28	100.8	7.5	1400	13	AAQ31436 Encodes a HeLa cel
29	100.8	7.5	1488	15	AAQ56910 DNA encoding a gly
30	100.8	7.5	2175	12	AAQ14382 Clone 7.2 encoding
31	100.8	7.5	2175	24	AAQ14383 Human cDNA clone 7
32	100.8	7.5	2861	12	AAQ14383 Clone 1 encoding 1
33	100.8	7.5	2861	21	AAF21133 Human low adenosin
34	100.8	7.5	2861	21	AAA35011 Human adenosine re
35	100.8	7.5	2861	21	AAQ17083 Human cDNA clone 1
36	100.8	7.5	3647	12	AAQ13333 GDP-Fuc-beta-D-gal
37	100.8	7.5	3647	15	AAQ56909 DNA encoding a gly
38	100.8	7.5	3647	18	AAT61678 Human alpha(1,3)-f
39	97.6	7.3	2134	18	AAT59506 Human myeloid deri
40	66.6	4.9	1578	23	ABL11867 Drosophila melanog
41	66.6	4.9	4229	23	ABL11866 C. elegans alpha-1
42	54.4	4.0	891	21	AAZ92670 Murine alpha-1,3-f
43	52.2	3.9	1353	21	AAZ51685 C. elegans alpha-1
44	49.8	3.7	90	21	AAC24928 Human secreted pro
45	44.8	3.3	2094	17	AAT33198 Alpha-1 A/D adrena

ALIGNMENTS

RESULT 1				
AAZ92647				
ID	AAZ92647	standard; cDNA; 2676 BP.		
XX				
AC	AAZ92647;			
XX				
DT	05-JUN-2000	(first entry)		
XX				
DE	Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.			
XX				
KW	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;			
KW	brain; kidney; recombinant expression; transgenic animal; knockout			
KW	mutant; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;			
KW	treatment; cancer; human; ss.			
XX				
OS	Homo sapiens.			
XX				
PH	Key	Location/Qualifiers		
FT	CDS	295..1374		
FT		/*tag= a		
XX		/product= "Human alpha-1,3-fucosyltransferase"		
PN	W0200006708-A1.			
XX				
PD	10-FEB-2000.			
XX				
PF	29-JUL-1999; 99WO-JP04092.			
XX				
PR	29-JUL-1998; 98JP-0213823.			
XX				
PA	(KIOW) KIOWA HAKKO KOGYO KK.			
XX				

PI Narimatsu H, Kudo T, Sasaki K;
XX
XX WPI: 2000-183120/16.
DR P-PSDB: AAY80996.
XX
XX Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer. -
XX
XX Claim 4; Page 143-150; 172pp; Japanese.

The invention relates to a novel alpha-1,3-fucosyltransferase which transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglucosamine (generating the Lewis x or y epitope). It does not transfer a fucose moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglucosamine and therefore does not generate the sialyl-Lewis x epitope. The invention also relates to DNA sequences encoding alpha-1,3-fucosyltransferase and expression vectors and host cells comprising these DNA sequences. The invention additionally encompasses the preparation of alpha-1,3-fucosyltransferase via the culture of transformed cells or by expression of the protein in a transgenic animal; antibodies which recognise alpha-1,3-fucosyltransferase; methods for screening potential inhibitors or potentiators of alpha-1,3-fucosyltransferase activity or expression; the preparation of compounds having fucose-containing sugar chains by use of the protein; and knockout non-human animals lacking alpha-1,3-fucosyltransferase. Alpha-1,3-fucosyltransferase has a similar substrate range to the known FUC-TV and is expressed mainly in brain and kidney tissues. Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies, potentiators and inhibitors may be used for the treatment and diagnosis of diseases of the brain and kidney, and of cancers. They may be used for the identification of substances which affect the activity or expression of alpha-1,3-fucosyltransferase, such substances may be used therapeutically. The knockout animals can be used to study the mechanisms of action and expression of alpha-1,3-fucosyltransferase. Sequences AA292646 and AA292647 represent cDNAs encoding human alpha-1,3-fucosyltransferase (AA280996).

Sequence 2676 BP; 819 A; 527 C; 476 G; 854 T; 0 other;

Query Match	100.08;	Score 1346;	DB 21;	Length 2676;
-------------	---------	-------------	--------	--------------

```
Matches 1346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ACTGCTCTCCCTGCAGCTCCCCGGCCCCCGCTGTGCTGCTCGCTCGGTGTCCCCCAG 60

Db 29 ACTGCTCTCCCTGCGAGCTCCCCCGCCCGCGCTGTGCGCTGCCCTGGGTGTCCCCCAG 88

61 CCCAGTCGGCTCTTAGGACAGCGCCGCCACCGCCCTGGCCCTGCCCTGCCCTCCCTGCG 12

db 89 CCCAGTCGGCTCTTAGGACAGCGCCGCCACCGCCCTGGCCCTGCCCTCTGCG 14

0Y 121 CCGCGAGCCCTCGCGAGCGCCCCGGATGCGCCTTACCCCTAGGACCGATTAGAAATGT 18

db 149 CCGCGAGCCCTCGCGAGCGCCCGGATGGCGTTACCCCTAGGACCGATTAGAACTG 20

181 AATACTCAAGGATTGATAATACAGTGAAGTATATACAACTGCTACGTCCTCCCA 24

Db 209 AATACCTCAAGGATTTGATATATACGTCAGTATACCACTGCTACGTCCTCCCA 26

[illegible][illegible]

201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

[illegible]

.....

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

421 ACTTCTTTCACCAAACTGATTATTATGAAGTACTATCTGGTGGTGGC 48

Db	449	ACTCTTTCCACCAAAAGTATTATTTAATGAACACTACTATTCGTGGTGGTGGC	508
Oy	481	CATTGGGACAGACCTTGACCTTACATCCCGGACAGCATATGTCACATCCAAAGATGCC	540
Db	509	CATTGGGCAAGACCTTGGACCTTACATCCCGGACAGCATATGTCACATCCAAAGATGCC	568
Oy	541	ATCCACACAGGAGACCTTGACCTTACATCCCGGACAGCATATGTCACATCCAAAGATGCC	600
Db	569	ATCCACACAGGAGACCTTGACCTTACATCCCGGACAGCATATGTCACATCCAAAGATGCC	628
Oy	601	ACATCAGTTGGGATCTGACAAATTTTACCTGAGCAAGCTAGGCCACCTTCCAGAAATGGA	660
Db	629	ACATCAGTTGGGATCTGACAAATTTTACCTGAGCAAGCTAGGCCACCTTCCAGAAATGGA	688
Oy	661	TTTGGATGAATTTGGAATVCAACCAACCTCACACTCCCAAAAAGATGGCATTTAGCACTTGT	720
Db	689	TTTGGATGAATTTGGAATVCAACCAACCTCACACTCCCAAAAAGATGGCATTTAGCACTTGT	748
Oy	721	TTAACTGACCTGACCTTACCGCGGATGATCAATATATCCAAAGTCCCTTATGGCTTCTTGA	780
Db	749	TTAACTGACCTGACCTTACCGCGGATGATCAATATATCCAAAGTCCCTTATGGCTTCTTGA	808
Oy	781	CGGTAAGACACAAATCCCTGTGTGTTTAAATGGCAAGCAAAAGAAATTTGGTGTGG	840
Db	809	CGGTAAGACACAAATCCCTGTGTGTTTAAATGGCAAGCAAAAGAAATTTGGTGTGG	868
Oy	841	TTTGAGTAAGCTGGAGACCCGAGCATGCCAGAGTCAGATTTTCAATVAGAGCTAAGCAAAA	900
Db	869	TTTGAGTAAGCTGGAGACCCGAGCATGCCAGAGTCAGATTTTCAATVAGAGCTAAGCAAAA	928
Oy	901	GCATTGAAATCCATACCTAGCGGCAACCATTTGAGAAATATGTCATGATTAATAAATTTGA	960
Db	929	GCATTGAAATCCATACCTAGCGGCAACCATTTGAGAAATATGTCATGATTAATAAATTTGA	988
Oy	961	TTTCCACCATATCTGCTTGTAAATTTTATCTTCCCTTGAAATTTCAATCCAAAGATTT	1020
Db	989	TTTCCACCATATCTGCTTGTAAATTTTATCTTCCCTTGAAATTTCAATCCAAAGATTT	1044
Oy	1021	ACATCAGGAAAGACTATACATGCTTTTCTGCTGGCTCTGTAAGTCTGTTCTTGAGAC	1080
Db	1049	ACATCAGGAAAGACTATACATGCTTTTCTGCTGGCTCTGTAAGTCTGTTCTTGAGAC	1100
Oy	1081	CATCTAGGAAAGACTATAGAGATTAATTCACAGCATTCATTCATCATGTGGAAGATT	1140
Db	1109	CATCTAGGAAAGACTATAGAGATTAATTCACAGCATTCATTCATCATGTGGAAGATT	1166
Oy	1141	ATAACTCTCCAGTAGACTATAGCAAGAATATCTGGAAGAGTGCACAAAAACATTAAGTAT	1200
Db	1169	ATAACTCTCCAGTAGACTATAGCAAGAATATCTGGAAGAGTGCACAAAAACATTAAGTAT	1228
Oy	1201	ACCTTAGTACTTTTAACTGAGAGAAAGATTTCACTGTAAATCTTCCACAGATTTTGGGAAT	1260
Db	1229	ACCTTAGTACTTTTAACTGAGAGAAAGATTTCACTGTAAATCTTCCACAGATTTTGGGAAT	1288
Oy	1261	CACATGCAATGTTTGGCTTGGCATATGTGAAAAGGATCAAGAAATTAAGTCTGTGGTA	1320
Db	1289	CACATGCAATGTTTGGCTTGGCATATGTGAAAAGGATCAAGAAATTAAGTCTGTGGTA	1346
Oy	1321	ATTTAGAGAAATGGTTTGGGAATTA 1346	
Db	1349	ATTTAGAGAAATGGTTTGGGAATTA 1374	
RESULT 2			
AAZ92646			
ID AAZ92646 standard; cDNA: 2822 BP.			
AC AAZ92646:			
XX	05-JUN-2000 (first entry)		
DT			
XX	Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:4.		

RESULT 2
AAZ92646
ID AAZ92646 standard; cDNA: 2822 BP

AC AAZ92646

DT 05-JUN-2000 (first entry)

Human alpha-1,3-fucosyltransferase cDNA, SEO ID NO:4.

XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
 KW treatment; cancer; human; ss.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1289..2368
 FT /tag="a
 FT /product="Human alpha-1,3-fucosyltransferase"
 XX
 PN WO200006708-A1.
 XX
 PD 10-FEB-2000.
 XX
 PD 29-JUL-1999; 99WO-JP04092.
 XX
 PR 29-JUL-1998; 98JP-0213823.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Narimatsu H, Kudo T, Sasaki K;
 XX WPI: 2000-183120/16.
 DR P-PSDB: AAY80996.
 XX
 PT Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT epitope and an antibody recognizing it useful for diagnosis of brain
 PT and kidney disease and cancer.
 XX
 PS Claim 4; Page 134-142; 172pp; Japanese.
 XX
 CC The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequences
 CC AA292646 and AA292647 represent cDNAs encoding human
 CC alpha-1,3-fucosyltransferase (AA292646).
 XX
 SQ Sequence 2822 BP; 918 A; 485 C; 477 G; 942 T; 0 other;
 XX
 Query Match 80.8%; Score 1088; DB 21; Length 2822;
 Best Local Similarity 100.0%; Pred. No. 1.5e-293;
 Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1341 TCCTGGGCTGTTTCATGCGATGCTTCTCATTTTACATCAAACTACCAACAGCTGATCT 1400
 QY 379 TCAGTCCAAATGGAATGAGCCAGCTCTGCTGTGAATGAAAACTTTTCCACCAAAA 438
 DB 1401 TCAGTCCAAATGGAATGAGCCAGCTCTGCTGTGAATGAAAACTTTTCCACCAAAA 1460
 QY 439 CTGATTTATTTTAAATGAACTACTATTCTGGTGTGGTGTGGCCATTTGGGACAGCTTTG 498
 DB 1461 CTGATTTATTTTAAATGAACTACTATTCTGGTGTGGTGTGGCCATTTGGGACAGCTTTG 1520
 QY 499 ACCCTACATCCGCGCAAGCAATGTCACATCCCAAGGATGCCATTCACACAGGACCGTT 558
 DB 1521 ACCCTACATCCGCGCAAGCAATGTCACATCCCAAGGATGCCATTCACACAGGACCGTT 1580
 QY 559 CACTGTACAACAATCCCATGTCAGTTCTGATCCATCAGCAGACATCAGTTGGGATCTGA 618
 DB 1581 CACTGTACAACAATCCCATGTCAGTTCTGATCCATCAGCAGACATCAGTTGGGATCTGA 1640
 QY 619 CAAATTTACCTCAGCAAGCTAGGCGACCTTCAGAAATGATTTGGATGAATTTGGAAT 678
 DB 1641 CAAATTTACCTCAGCAAGCTAGGCGACCTTCAGAAATGATTTGGATGAATTTGGAAT 1700
 QY 679 CACCACTCAGACTCCCGCAAAAGAGTGGCATGTAGACATCTTTTAACTGACTCTGACTT 738
 DB 1701 CACCACTCAGACTCCCGCAAAAGAGTGGCATGTAGACATCTTTTAACTGACTCTGACTT 1760
 QY 739 ACCGCGGTATTCAGATATTCAGATGCTGCTTATGCTTTCAGAGTAAAGCAATCCCT 798
 DB 1761 ACCGCGGTATTCAGATATTCAGATGCTGCTTATGCTTTCAGAGTAAAGCAATCCCT 1820
 QY 799 TCGTGTGTAAGTCCCAAGCAAAAGAGAAATTTGGTGTGCTGCTGTGTAAGTAACTGGAAAC 858
 DB 1821 TCGTGTGTAAGTCCCAAGCAAAAGAGAAATTTGGTGTGCTGCTGTGTAAGTAACTGGAAAC 1880
 QY 859 CTGAGATCCAGAGTCAAGTATATCAATGAGCTAGCAAAAGCAATTCATCACT 918
 DB 1881 CTGAGATCCAGAGTCAAGTATATCAATGAGCTAGCAAAAGCAATTCATCACT 1940
 QY 919 ACGGCAACCATTTTGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 978
 DB 1941 ACGGCAACCATTTTGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 2000
 QY 979 GTAAATTTATCTTCTTCTTGAATTCATCCAGAGATTAATCAATCAGGAAAGCTAT 1038
 DB 2001 GTAAATTTATCTTCTTCTTGAATTCATCCAGAGATTAATCAATCAGGAAAGCTAT 2060
 QY 1039 ACAATGCTTTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098
 DB 2061 ACAATGCTTTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120
 QY 1099 AGAATTTATTTCCAGAGATTTCAATTCATGTCAGTGAAGATTAATCAATCAGGAGAGC 1158
 DB 2121 AGAATTTATTTCCAGAGATTTCAATTCATGTCAGTGAAGATTAATCAATCAGGAGAGC 2180
 QY 1159 TAGCAAAATGATCTGAAGAGTGCAGAAAAACAATAGTTATCTTACTTTTAACT 1218
 DB 2181 TAGCAAAATGATCTGAAGAGTGCAGAAAAACAATAGTTATCTTACTTTTAACT 2240
 QY 1219 GAGAGGAAGATTTCACTGTAAATCTTCCACAGATTTTGGGATGACATGCTTTGGCTT 1278
 DB 2241 GAGAGGAAGATTTCACTGTAAATCTTCCACAGATTTTGGGATGACATGCTTTGGCTT 2300
 QY 1279 GCGATCATGTGAAGAGCAATCAAGATTAAGCTGTGTATTTAGAGAAATGCTTTT 1338
 DB 2301 GCGATCATGTGAAGAGCAATCAAGATTAAGCTGTGTATTTAGAGAAATGCTTTT 2360
 QY 1339 GGAATTTAA 1346
 DB 2361 GGAATTTAA 2368
 RESULT 3
 AA292645

XX	AA292645 standard; cDNA; 2170 BP
AC	AA292645;
DT	05-JUN-2000 (first entry)
DE	Murine alpha-1,3-fucosyltransferase cDNA.
XX	
XX	Alpha-1,3-fucosyltransferase: fucose; glycosylation; Lewis epitope;
KW	brain; kidney; recombinant expression; transgenic animal; knockout
KW	animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
KW	treatment; cancer; murine; mouse; ss.
XX	
OS	Mus sp.
XX	
FT	Key
FT	Location/Qualifiers
FT	115..1194
FT	/tag a
FT	/product "Murine alpha-1,3-fucosyltransferase"
XX	W0200006708-A1.
XX	
PD	10-FEB-2000.
XX	
PD	29-JUL-1999; 99MO-JP04092.
XX	
PR	29-JUL-1998; 98JP-0213823.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	
PI	Narimatsu H, Kudo T, Sasaki K;
DR	WPI: 2000-183120/16.
DR	P-PSDB; AAY80995.
XX	
PT	Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT	epitope and an antibody recognizing it useful for diagnosis of brain
PT	and kidney disease and cancer.
XX	
PS	Claim 4; Page 127-134; 172pp; Japanese.
XX	
CC	The invention relates to a novel alpha-1,3-fucosyltransferase which
CC	transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
CC	(generating the Lewis x or y epitope). It does not transfer a fucose
CC	moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
CC	and therefore does not generate the sialyl-Lewis x epitope. The
CC	invention also relates to DNA sequences encoding alpha-1,3-
CC	fucosyltransferase and expression vectors and host cells comprising
CC	these DNA sequences. The invention additionally encompasses the
CC	preparation of alpha-1,3-fucosyltransferase via the culture of
CC	transformed cells or by expression of the protein in a transgenic animal;
CC	antibodies which recognise alpha-1,3-fucosyltransferase; methods for
CC	screening potential inhibitors or potentiators of
CC	alpha-1,3-fucosyltransferase activity or expression; the preparation of
CC	compounds having fucose-containing sugar chains by use of the protein;
CC	and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
CC	Alpha-1,3-fucosyltransferase has a smaller substrate range to the known
CC	FUC-TIV and is expressed mainly in brain and kidney tissues.
CC	Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
CC	potentiators and inhibitors may be used for the treatment and diagnosis
CC	of diseases of the brain and kidney, and of cancers. They may be used for
CC	the identification of substances which affect the activity or expression
CC	of alpha-1,3-fucosyltransferase; such substances may be used
CC	therapeutically. The knockout animals can be used to study the mechanisms
CC	of action and expression of alpha-1,3-fucosyltransferase. Sequence
CC	AA892645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
CC	(AA890995), and sequences AA892670-892673 are murine
CC	alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
CC	amplification of cDNA ends).
XX	
SO	Sequence 2170 BP; 715 A; 395 C; 368 G; 692 T; 0 other:

Query Match 73.3%; Score 986.8; DB 21; Length 2170;

Best Local Similarity 90.0%; Pred. No. 2.6e-265;
Matches 1069; Conservative 0; Mismatches 117; Indels 2; Gaps 1;

OY	159	CCAGAGACGGATTTGAAATGTAATATACACAGATTTGATATATATACAGGAAGTATGTA	218
Db	9	CGCGGGATGCGCTTTAGAAATGTATATACACAGCATCTGTATAGCACCGTGGAGTATGTTCA	68
OY	219	ACAACTGTCTACGTGCTTCCCATGATATGTCTCTATATTGAAAAATATGTGATCAACA	278
Db	69	GCATCATCTTC--TGGTTTCCATGATGATGTCTCTACCTGTAAGAAATATATGATCAACAA	126
OY	279	TCCAAGGAATTCCTGCGCCATTTTAAATGTCTGCATTAATCTGGGCTGTTCAATGGCA	338
Db	127	TCCAAAGGCAATTCCTGCGCCATTTTAAATGTCTGCATTAATCTGGGCTGTTCAATGGCA	186
OY	339	TGTCTTCTCATTTACATCAAAACCTTACCAACAGCTGATCTTCACTCCATGAAATGACCC	398
Db	187	TGTCTTCTCATTTATATCAAGCCCAACACAGCTGGGCTGTTCACTCCATGAAATGACCC	246
OY	399	AGCTCTGCGTGAATAAGAAAACTCTTTCCACCAAAAGCATTAATTTAATGAACCT	458
Db	247	AGTCTGTGCTGAATAAGAAAAATTTCTTCTCCACAAAAGCATTAATTTAATGAACCT	306
OY	459	ACTATCTGGTGGGTGGGTGGCCATTTGGGACACCTTTGACCTTACATCTCGGCCAACCA	518
Db	307	ACCATCTGGTGGGTGGGTGGCCATTTGGGACACCTTTGACCTTACATCTCGGCCAACCA	366
OY	519	ATGTTTCAATCCAAAGATATGCCATCTCACACAGGACCGCTTCACTGTACAAACAATCCAT	578
Db	367	ATGTTTCAATTCCAAGGATGTCATCTCACACAGGACCGCTTCACTGTACAAACAATCCAT	426
OY	579	GCAGTTCGATCCATCAGACGAGACATCGTTGGGATCTGACAAATTTACCTCAGCAACT	638
Db	427	GCGGTCTGTATTCACCTTAAGACATCAGCTGGGATCTGACAACTTACCTCAGCAAGCC	486
OY	639	AGGCGACCCCTCCAGAAATGATTTGGATGTAATTTGGAATACCAACTACACTCCCA	698
Db	487	AGGCGACCCCTCCAGAAATGATTTGGATGTAATTTAGACTACCAACTACACTCCCA	546
OY	699	AAGAGTGGCATGTGACACTGTTTAACTGACTCTACTTACCGCGTATTCAGATATTC	758
Db	547	AAGAGTGGCATGTGACACTGTTTAACTGACTCTACTTACCGCGTATTCAGATATTC	606
OY	759	CAATGGCTTATGGCTCTCTGAGGGTAAGACAAATCCCTGCTGTTGTAAGTGGCAAGC	818
Db	607	CAATGGCTTATGGCTCTCTGAGGGTAAGACAAATCCCTGCTGTTGTAAGTGGCAAGC	666
OY	819	AAAGAGAAATTTGGTCTGCTGGTGTGTAGTACTGTGAACCCCTGAGCATCCGAGTCAAG	878
Db	667	AAAGAGAAATTTGGTCTGCTGGTGTGTAGTACTGTGAACCCCTGAGCATCCGAGTCAAG	726
OY	879	TATTACATGAGCTTACGAAAAAGCATTTGAATTCATACCTACGGGCAAGCATTTGAGAA	938
Db	727	TATTACAAACGAGCTACGAAAGATATTTGAATTCACACCTATAGGCCCAACATTTGGAGAA	786
OY	939	TATGTCAATGATATAAATTTGATTTCCATCCATATCTGCTGTGAATTTATCTTCTCTTT	998
Db	787	TATGTCAACGATATAAATTCGATTTCCACCATATCTGATGTGAATTTATCTTCTCATTT	846
OY	999	GAAATTCATATCCAAAGATATTCATCAGCGAAACCTTTACAAATGCTTTTGTGCTGGC	1058
Db	847	GAAATTCATATTCAAAGATATTCATCAGAAACCTCTACAAATGCAATTTTGTGCTGGT	906
OY	1059	TCTGTACCTGTGTTCTTGGGACCATCTAGGGAAGAACTATGAGAAATATTTCCACAGAT	1118
Db	907	TGATGATCTGTGTGTCCTGGGTCCATCTAGGGAAGAACTATGAGAAATATTTCCACATGAT	966
OY	1119	TCATTTCAATCTGTGGAAAGATATTAATCTTCCAGTACCTTACCAAGATATCTGAAAGAA	1178
Db	967	TCATTCATCTGTGGGAAGATTTTAATCTTCCAGTGAAGTATACCAAAATCTGGAAGAA	1028
OY	1179	GTGCACAAAAACATTAAGTTATACCTAGTTACTTTTAACTGAGAGAGATTCACCTGA	1238

DB 1027 GTTGACAAAACATAAGTGTACCTTACTTACTGAGAAAGCATTTTACTGTA 1086
QY 1239 AATCTCCACGATTTTGGGATCAGATGATGTTGGCTTGCATCATGTGAAAAGCAT 1298
DB 1087 AACCTACACAGGTTTGGGATCAGATGATGTTGGCTTGCATCATGTGAAAAGCAT 1146
QY 1299 CAGGATATTAAGCTGTGTGTAATTTAGGAAATGTTTGGGATTTAA 1346
DB 1147 CAGGATATTAAGCTGTGTGTAATTTAGGAAATGTTTGGGATTTAA 1194

RESULT 4
AAZ92672
ID AAZ92672 standard; cDNA; 2036 BP.
XX AAZ92672;
XX
XX
XX 05-JUN-2000 (first entry)
DE Murine alpha-1,3-fucosyltransferase gene exon 1 DNA, SEQ ID NO:30.
XX
XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
KM brain; kidney; recombinant expression; transgenic animal; knockout
KM animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;
KM treatment; cancer; murine; mouse; ss.
XX
XX Mus sp.
OS
XX WO200006708-A1.
XX
XX 10-FEB-2000.
PD
XX 29-JUL-1999; 99WO-JP04092.
PF
XX 29-JUL-1998; 98JP-0213823.
PR
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Narimatsu H, Kudo T, Sasaki K;
PI
XX WPI; 2000-183120/16.
XX
XX Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
PT epitope and an antibody recognizing it useful for diagnosis of brain
PT and kidney disease and cancer.
XX
XX
XX Example 8; Page 162-164; 172pp; Japanese.
XX
XX The invention relates to a novel alpha-1,3-fucosyltransferase which
CC transfers a fucose moiety to galactosyl-beta-1-4-N-acetylglucosamine
CC (generating the Lewis x or y epitope). It does not transfer a fucose
CC moiety to alpha-2,3-sialyl-galactosyl-beta-1-4-N-acetylglucosamine
CC and therefore does not generate the sialyl-Lewis x epitope. The
CC invention also relates to DNA sequences encoding alpha-1,3-
CC fucosyltransferase and expression vectors and host cells comprising
CC these DNA sequences. The invention additionally encompasses the
CC preparation of alpha-1,3-fucosyltransferase via the culture of
CC transformed cells or by expression of the protein in a transgenic animal;
CC antibodies which recognize alpha-1,3-fucosyltransferase; methods for
CC screening potential inhibitors or potentiators of
CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
CC compounds having fucose-containing sugar chains by use of the protein;
CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
CC FUC-TIV and is expressed mainly in brain and kidney tissues.
CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
CC potentiators and inhibitors may be used for the treatment and diagnosis
CC of diseases of the brain and kidney, and of cancers. They may be used for
CC the identification of substances which affect the activity or expression
CC of alpha-1,3-fucosyltransferase; such substances may be used
CC therapeutically. The knockout animals can be used to study the mechanisms
CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
CC AAZ92645 represents cDNA encoding murine alpha-1,3-fucosyltransferase

CC (AA180995), and sequences AAZ92670-292673 are murine
CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
CC amplification of cDNA ends).
XX
SQ Sequence 2036 BP; 666 A; 365 C; 348 G; 657 T; 0 other:
Query Match 69.8%; Score 939.2; DB 21; Length 2036;
Best Local Similarity 91.5%; Pred. No. 5,1e-252;
Matches 995; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 259 GAAAAATTTATGATCAATCAATCCAAAGAAATTTTGGCCATTTTAAATGCTGCATTA 318
DB 1 GAAAAATTTATGATCAATCAATCCAAAGAAATTTTGGCCATTTTAAATGCTGCATTA 60
QY 319 TCTGGGCTGTTCATGAGCATGCTCTCATTTACATCAACCTACCAAGCTGCATCT 378
DB 61 TCTGGGCTGTTCATGAGCATGCTCTCATTTATTCAGGCCACCAAGCTGCATCT 120
QY 379 TCAGTCCAAATGGAATGAGCCAGCTGTGTCTGTAATAAATGAAAACTTTTCCACCAAA 438
DB 121 TCAGTCCAAATGGAATGAGCTGTGTCTGTAATAAATGAAAACTTTTCCACCAAA 180
QY 439 CTGATATTTTATGAACTACTATCTGTGGGTGTGGCATTTGGGACAGCTTTG 498
DB 181 CTGATATTTTATGAACTACTATCTGTGGGTGTGGCATTTGGGACAGCTTTG 240
QY 499 ACCTACATCTGTCGAAGCAATGTTCAACATCCAGATGCCATTCACAGGACGCT 558
DB 241 ACCTACATCTGTCGAAGCAATGTTCAACATCCAGATGCCATTCACAGGACGCT 300
QY 559 CACTGTACAAACAAATCCATGCAATTCATCCATCCAGGACGATCGGATCTGA 618
DB 301 CACTGTACAAACAAATCCATGCAATTCATCCATCCAGGACGATCGGATCTGA 360
QY 619 CAAATTTACCTGAGCAAGGTAGGCGCCCTTCCAGAAATGATTTGATTTGAAAT 678
DB 361 CAAATTTACCTGAGCAAGGTAGGCGCCCTTCCAGAAATGATTTGATTTGAAAT 420
QY 679 CACCAACTCACAATCCCAAAAGAGGATGAGCATGAGCATTTGTTAACTGACTGCT 738
DB 421 CACCAACTCACAATCCCAAAAGAGGATGAGCATGAGCATTTGTTAACTGACTGCT 480
QY 739 ACCGCGGTATTCAGATATTCAGAGTCCCTTATGCTTCTTGACGGTAAACAAATCCCT 798
DB 481 ACCGCGGTATTCAGATATTCAGAGTCCCTTATGCTTCTTGACGGTAAACAAATCCCT 540
QY 799 TCGTGTTTAAGTCCCAACCAAGAAATGCTGTGCTGGGTGTGATTAAGTGAAC 858
DB 541 TCGTGTTTAAGTCCCAACCAAGAAATGCTGTGCTGGGTGTGATTAAGTGAAC 600
QY 859 CTGAGCATGCCAGAGTCAAGTATTCATGAGTGAAGCAAAAGCATTTGAAATCCATACCT 918
DB 601 CTGAGCATGCCAGAGTCAAGTATTCATGAGTGAAGCAAAAGCATTTGAAATCCATACCT 660
QY 919 ACGGCAAGCAATTTGGAGATATGTCATGATTAATAAATTTGATTCCTACATATCTGCT 978
DB 661 ACGGCAAGCAATTTGGAGATATGTCATGATTAATAAATTTGATTCCTACATATCTGCT 720
QY 979 GTAAATTTATCTTCTTCTTGAATAATTCACAGATTTACATCAGGAAAGCTAT 1038
DB 721 GTAAATTTATCTTCTTCTTGAATAATTCACAGATTTACATCAGGAAAGCTAT 780
QY 1039 ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098
DB 781 ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 1099 ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158
DB 841 ACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 1159 TAGCAAAATATCTGAAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 1218
DB 901 TAGCAAAATATCTGAAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 960

QY	1219	GGAGGAAGAAATTCACGTCTAAATCTTCACGATTTGGGAATCACATGATGTTGGCTT	1278
	961	GGGAAAGGATTTACTGTAAACCTTCACAGGTTTTGGGAATCACATGATGCCCGCAT	1020
Db			
QY	1279	CGGATCTATGTGAAAAAGGCACGAATATTAAGCTCTGGTAATTGAGAAATGGTTTT	1338
Db	1021	CGCGACCTGTAAAAAGGCATCAAGAAATATTAAGTCTGTGGTAATTGAGAAATGGTTTT	1080
QY	1339	GGAAATTAA	1346
Db	1081	GGAAATTAA	1088

RESULT 5
AAT59505
ID AAT59505 standard; cDNA; 1814 BP.
XX
AC AAT59505;
XX
DT 06-MAY-1997 (first entry)
XX
DE Murine myeloid-1 lineage alpha-(1,3)-fucosyltransferase cDNA.
XX
KW Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM,
XX septic shock; septicemia; therapy; ss.
XX
US Mus sp.

FT	Key	Location/Qualifiers
FT	CDS	325.1353
FT		/*tag- a
XX		
XX	W09640881-AL.	
XX		
PD	19-DEC-1996.	
XX		
PF	08-MAY-1996;	96W0-0506427.
XX		
PR	07-JUN-1995;	9505-0483151.

XX (Geno) GEN HOSPITAL CONF
XX
PI Holgersson J, Seed B;
XX
DR WPI; 1997-108639/10.
DR P-PSDB; AA111820.

PT New murine alpha-(1,3)-fucosyltransferase - for fucosylating an
PT antibody to protect mammals against e.g. septic shock or septicæmia
XX
PS Claim 6; Fig 3; 58pp; English.

A CDN clone (AA759505) codes for murine myeloid lineage alpha-(1,3)-fucosyltransferase (AAW11820), an enzyme which has a relatively strict substrate requirement for sialylated N-acetylactosamine, which can account for the presence of the sialyl-Le^x epitope on murine cells, and which is more effective than Fuc-IV in support of E-selectin-mediated COS cell adhesion. It was isolated from a murine myeloid 32D c13 cell cDNA library by identification of a clone capable of directing the expression of sialyl-Le^x determinants. Transformed host cells (e.g. 32D c13 or human 293 cells) expressing the alpha-(1,3)-fucosyltransferase, and pref. another fucosyltransferase such as human Fuc-IV (see also AA759506), can be used to fucosylate an antibody (e.g. IgG or IgM) for use in protecting an animal against an adverse immune reaction, esp. septic shock or septicaemia.

Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other;

Query Match	9.58;	Score 127.4;	DB 18;	Length 1814;
Best Local Similarity	52.18;	Pred. No. 4.2e-25;		
Matches 334;	Conservative	0;	Mismatches 301;	Indels 6;
				Gaps 2

OY	535	GATGCACTGTCACAGGAGCGCTTCTACTGTATACAAACAAATCCCATCTGATTCATCCATC	594
Db	548	GCATCCCTGTGATGTCGTAAACCGGAGGCTGTATGACCAAGTCTGATGCTGTGGTCTTCCACC	607
OY	595	ACCGAGACATCATGTTGGGATCTGACAAATTTTACCTCAGACACCTGAGCCACCTTCACGA	654
Db	608	ACCTGTAGCTGTCAAACCCCGCAATCTCTCTCTACCTGTGACCAAGGCCACACAGCACAC	667
OY	655	AATGATTTGGATNGAATTTGGATTCACCAACTCACACTCCCCAAMAGATGSCATTTGAGC	714
Db	668	CTTGGGCTGTGGGCTCCATGGAATTCGCCACAGTAATGCCATGCTCTCCATGCTCTCCGGG	727
OY	715	ACTGTTTAAACCTGACCTGCTACTTACCGCCGCTGATTCAGATATCCAGTGCCTTAAGCT	774
Db	728	GCATCTTCAACTGTGGTGTCTAGCTATTCGGGTGATTCATTTTGTATACCTCTACGGTTC	787
OY	775	TCTTGACGATAGACACAATCCCTGTGTGTGTTGAATGCGCAAGCAAGAATATGGGT	834
Db	788	GCTTGGAGCCCTCTCTGCGGCC---CACATGCCCATCTACCGGCCAAMAGCAGATGGCTG	844
OY	835	GCTGGGTGTGAGTAATCTGGAACCCCTGAGATCCAGATCCAGATCAAGTAATTCATATGAGTAA	894
Db	845	CTTGGTGTGATCAGCAATTTTCCAGGAGCGCAGCAGCGTGTCAAAGCTGTACGGCAGCTGG	904
OY	895	GCAAAAGCATTTGAATCCATACCTACGGGCGACGATTTGGAGATATGTCAATGATATAAA	954
Db	905	CCCCCATCTGTCAAGTGGATGTGTGGTGCGCCGACGGGACGGCCCCATGATGCTTAATTT	964
OY	955	ATTGATTTTCATACATATCATGCTTGTGTAAATTTTATCTTTCTTGAATAATCAATCACA	1014
Db	965	GTCGTGCGCCCATCTTTTGGCCCGGACCGCTTCTACCTGCGCTTTGGAACCTACAGCATTC	1024
OY	1015	AGGATTTACATCAGCAAAACGTAT---ACATATCTTTTGTGGCTGTGTACCTGTGTG	1071
Db	1025	GGGACTACATACGTAGAAAGTTCGTGGCGCAATCCCTGGGGGGGTGGTGTGTACCCGTGG	1084
OY	1072	TTTTCGGACCATTTAGGGAAGAAACATATGAGATTAATTTTCACAGATTTATTCATTCATTCAG	1133
Db	1085	CGCTGTGAGACCTCTCTCGGGCCACCTACAGAGGCTTTTGTGCACACAGATGCTTTGTACACG	1144
OY	1132	TGCAAGATTATACTCTTCCAGTGAAGTGTAGCAAAAGTATCTG	1172
Db	1145	TGGAAGACTTATAGCTCTGCCCGGTAACTGTGGCTGTCTTCTCG	1185

RESULT 6	
AAT85219	
ID	AAT85219 standard; DNA; 3594 BP
XX	
AC	AAT85219;
XX	
DT	02-MAR-1998 (first entry)

Mouse alpha-fucosyltransferase Fuc-TVII gene.

KW Alpha-fucosyltransferase; Fuc-TVII gene; mouse;
KW sialyl Lewis x determinant; oligosaccharide; antiinflammatory;
KW inhibitor; ds.

OS Mus musculus.

FH	Key	Location/Qualifiers
EE	CD	006 3083

```
/*tag= a
/nota= "includes introns"
```

ET
ET
exon

intron
ET
ET

32

FT exon 670...1149

```

FT      /tag= d
FT      /number= 2
FT      /codon_start= 966..998
FT      Intron
FT      /tag= e
FT      /number= 2
FT      /number= 2
FT      /tag= f
FT      /number= 3
FT      /codon_start= 1947..1949
FT      /note= "alternative ATG start codon"
FT      Intron
FT      /tag= g
FT      /number= 3
FT      /number= 3
FT      /tag= f
FT      /number= 4
FT      /codon_start= 2126..2128
FT      /note= "alternative ATG start codon"
XX
XX      M09732889-A1.
XX      12-SEP-1997.
XX
XX      07-MAR-1997; 97MO-US03184.
XX
XX      08-MAR-1996; 96US-0613098.
XX
XX      (UNMI ) UNIV MICHIGAN.
XX
XX      Gersten KM, Lowe JB, Natsuka S;
XX
XX      WPI: 1997-47053/43.
XX      P-PSDB: AAW27138 and AAW26671.
XX
XX      Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g.
XX      screening for inhibitors used as antiinflammatory agents
XX
XX      Claim 1: Fig 2; 91pp; English.
XX
XX      This DNA sequence comprises the mouse Fuc-TVIII gene encoding
XX      alpha-fucosyltransferase. It was isolated from a 3T3 cell genomic
XX      library by screening with a probe from the catalytic domain of the
XX      human Lewis alpha(1,3/4)fucosyltransferase (Fuc-TVIII).
XX      Examination of the mouse Fuc-TVIII locus identified 3 Met codons
XX      that may initiate translation of alpha(1,3) fucosyltransferases
XX      with different cytosolic domains encoded by exons 2 and/or 3, but
XX      with identical Golgi-localised catalytic domains encoded by exon 4
XX      (see AAW27138 and AAW26671). The proteins and nucleic acids can
XX      be used to construct animal cell lines able to post-translationally
XX      modify oligosaccharides on cell surface, intracellular or secreted
XX      proteins and lipids, to isolate reagents for efficient enzymatic
XX      production of oligosaccharides, to generate antibodies to
XX      glycosyltransferase inhibitors useful as diagnostic reagents, to screen for
XX      fucosyltransferase inhibitors and inactivators, especially those
XX      that act as antiinflammatory agents, for genotyping individuals at
XX      the fucosyltransferase locus and for in vitro synthesis of sialyl
XX      Lewis x tetrasaccharide.
XX
XX      Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other;
SQ
Query Match          9.48; Score 127; DB 18; Length 3594;
Best Local Similarity 52.3%; Pred. No. 7.9e-25;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

```

```

QY      655 AATGATTTGATGATTTGATATCCACTCACTCCCAAAAGAGGCAATGAGC 714
D      2397 CTGGGCTCTGGGCTCCATGGAATGCCAGATATACCATGATGCTTCGCGG 2456
QY      715 ACTGTTTAACCTGACCTGACCTGACCGCGGTGATGATATCCAGTGGCTATGGCT 774
D      2457 GCATCTTAACCTGGGTGCTGAGCTATGCGCGTATGATATCTTGTACCTACGCTC 2516
QY      775 TCTTGACGGTAAGACAAATCCCTGCTGTGTTGAAGTCCCAAGCAAGAAATGGGT 834
D      2517 GCTTGAGACCTCTCTCTGGGCC---CACATCCCATCTGCGGCCCAAAAGAGATGGCTG 2573
QY      835 GCTGGGTTGTGATGATGTAACCTGGAACCTGAGCATGCCAGATCACTATTACATAGCTAA 894
D      2574 CCTGGGTGATCAGCAATTTCCAGAGGCGCAGCAGCGTGCAGAAAGTGAACCGGAGCTGG 2633
QY      895 GCAAAAGCATTGAAATCCATACCTACGAGGAGCAAGATTTGGAGAAATATGATATAAA 954
D      2634 CCCCTCATCTCCAGGTGATGTGTGCGGCCGACGCGAGCGCCCTATGCGCTAATTT 2693
QY      955 ATTTGATTCCTACATATCTGCTGTGAATTTTAATCTTCTTTGAAATTCATTCACA 1014
D      2694 GTCTGCTCCACACTTTGGCCCGGATACCGCTCTCTACCTGCTTTGAGAACTACAGATC 2753
QY      1015 AGCATTTACATACGGAAGACTAT--ACAATGCTTTTGTGCGCTGCTGATCCTGTTG 1071
D      2754 GGGACTACATCATCAGAGAAATTTGCGCGCAATGCCCTGGCGGTGTGCTATACCGCTG 2813
QY      1072 TTCTGGGACCATCAGGAAACATATGAAATTTATTTCCAGAGATTCATTCATG 1131
D      2814 CGTGGGACCTCTCTGGGCGCACCTACGAGGCTTTTGTGCGCACAGATGCCCTTTGACACG 2873
QY      1132 TGGAAATTATTACTCTCCAGTAGCTAGC 1162
D      2874 TGGACACTTCAGCTCTGCCGTGACTGCG 2904

RESULT 7
AAQ77732
ID AAQ77732 standard; cDNA to mRNA; 1701 BP.
XX
AC      AAQ77732;
XX
DT      22-JUN-1995 (first entry)
XX
DE      Human alpha-1,3-fucosyltransferase cDNA.
XX
KW      Alpha-1,3-fucosyltransferase; sialylated Lewis blood group antigen;
KW      Lewis x; Lewis a; glycosylation; glycoprotein; glycolipid; ds.
XX
OS      Homo sapiens (THP-1 cells).
XX
FH      Key Location/Qualifiers
FT      CDS 123..1151
FT      /tag= a
FT      /product= alpha-1,3-fucosyltransferase
PN      W09423021-A.
PD      13-OCT-1994.
XX
XX      28-MAR-1994; 94WO-JP00496.
XX
XX      29-MAR-1993; 93JP-0069016.
XX
PA      (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI      Hanai N, Kurata K, Nishi T, Sasaki K;
XX
XX      WPI: 1994-333183/41.
XX      P-PSDB: AAR63215.
XX
XX      Alpha-1,3-fucosyl-transferase DNA and protein - useful for prodn.
PT

```


PT of active sugar chains, e.g. sialylated Lewis X
 XX
 XX Claim 3; Page 99-102; 155pp; Japanese.
 PS
 CC The sequence coding for alpha-1,3-fucosyltransferase was isolated
 CC from a human THP-1 cell cDNA library. The enzyme catalyzes the
 CC addition of fucose via a 1,3-linkage to N-acetylglucosamine
 CC contained in glycoproteins and glycolipids, e.g. for sialylation of
 CC Lewis blood group antigens.
 XX

SQ Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;

Query Match 8.5%; Score 114.4; DB 15; Length 1701;
 Best Local Similarity 50.8%; Pred. No. 1.8e-21;
 Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

OY 522 TTCACATCCAGAGATGCCATCTGCACAAAGCGCTGATACAAACAAATCCCATGCA 581
 DB 333 TACGGCATGCCCGCTGCTGCCACTGATGTCACCAACGACCTGCTGCGCCAGCCGACGCC 392
 OY 582 GTTCTGATCATCACCAGACATTCAGTTGGATCTGACAAATTTACTCAGCAAGCTAGG 641
 DB 393 GTGGTCTTCACACACGCGAGCTGCAGAACCGCGGCTCCACCTGCTGCGCCAGGG 452
 OY 642 CCACCTTCCAGAAATGATTTGATGAAATTTGAAATCAGCACTACCTCCCAAAAG 701
 DB 453 CCGCGAGGACGACCCCTGGGTGTGGGCTCCATGAGTCTCCTAGCCACACCCAGCGCTC 512
 OY 702 AGTGCAATGAGCACTGTGTTAACTGACTGACTGACCGCGCTGATTCAGATATCCAA 761
 DB 513 AGCCACCTCCAGGAGCAATCTCAACTGGGCTGAGCTACCGCGGCTGCGACATCTTT 572
 OY 762 GTGCTTATGCTTCTTGACGGTAAACCAATCCCTGCTGCTTGAAGTCCCAAGCAA 821
 DB 573 GTCCCTATGCTGCTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 629
 OY 822 GAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
 DB 630 AGGAGGTGCGCGCTGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
 OY 882 TACAAATGAGCTAAGCAAAAGCATTTGAATCCATACCTGAGGCGCAAGCTTTGGAAAT 941
 DB 690 TACCGGACCTGCGCCCTCATCTGCGGGTGAATGCTTTGGCGCCCAATGGAGCGCA 749
 OY 942 GTCAATGATAAATTTGATTTCTTCAATCTGCTTTGTAATTTTATCTTTCTTGA 1001
 DB 750 CTGTGCGCCAGTGCCTGTGCTGCCACCGGTGGCCAGTACCGCTTCTTACCTGCTT 809
 OY 1002 AATTCAATCCAGAGATTTACATCACGAGAAAGCTAT--ACAATGCTTTTCTGCTGGC 1058
 DB 810 AACTCTACAGCAGCGGACTACATTTAGCGAATTTCTGCGCAACGCACTGCTGGCTGGC 869
 OY 1059 TCTGACTGTTGTTCTGAGGACATCTAGGAAACATGATGAAATTTATATTCAGAGAT 1118
 DB 870 ACGTGTCCAGTGTGTGGGGCCCGCCAGCGGCGCACTATGAGGCGCTTGTGCTGCGG 929
 OY 1119 TCATTCATCATGTGGAAGATTAATACTCCCAATGAGTGCAGC 1162
 DB 930 GCCTTCGTGATGTGATGACTTTGGCTCAGCGCAAGAGCTGCG 973

RESULT 8
 AAF21136
 ID AAF21136 standard; DNA: 1701 BP.

AC AAF21136;

DT 14-MAR-2001 (first entry)

DE Human low adenosine antisense oligonucleotide related sequence #2703.

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

KW human; airway disorder; bronchoconstriction; lung inflammation;

KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antisthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 PN WO20062736-A2.
 PD 26-OCT-2000.
 XX
 XX 24-MAR-2000; 2000WO-US08020.
 PF 06-APR-1999; 99US-0127958.
 XX
 XX (UYEC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-679539/66.
 XX
 PT Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -
 PS
 XX
 XX Disclosure; Page 1034; 1592pp; English.
 XX
 CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a "universal" or alternative base.
 CC (II) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antisthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and/or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulin and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokines and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC transmitters, CNS and peripheral nervous and non-nervous system
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergic asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 1701 BP; 290 A; 527 C; 589 G; 295 T; 0 other;

Query Match 8.5%; Score 114.4; DB 21; Length 1701;
 Best Local Similarity 50.8%; Pred. No. 1.8e-21;
 Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

OY 522 TTCACATCCAGAGATGCCATCTGCACAAAGCGCTGATACAAACAAATCCCATGCA 581
 DB 333 TACGGCATGCCCGCTGCTGCCACTGATGTCACCAACGACCTGCTGCGCCAGCCGACGCC 392
 OY 582 GTTCTGATCATCACCAGACATTCAGTTGGATGTGACAAATTTACTCAGCAAGCTAGG 641

Db	393	GTGGCTTCCACACACCGCAGCTGCAGAACCCCGGCGGTCCACACTGGCCCTGGCCACGAG	452
OY	642	CCACCCCTTCCAGAAATGGATTGGATTTGGATTGGATTACACCACTCAGCACTCCCCAAAAG	701
Db	453	CCGGAGGAGGAGCGCTGGGGGTGGGGCTCCATGAGATCTCTCAGCCACACCCAGGACCTC	512
OY	702	AGTGGCAATTAAGCACTTGTTTAACCTGACCTGACTTACCGCGGTGATTCAGATATCCAA	761
Db	513	AGCCACCTCCGAGGACATCTTCACTCAGCTGGGTGCTAGCTACCGCGGACCTCGGACATCTTT	572
OY	762	GTGCTTATATGGCTTCTTTCAGCGGTAAACACAAATCCCTTGTGTTTGAAGTGCACAA	821
Db	573	GTGCGCTTATATGGCGCTGGGAGCCCACTGGGGGCGCT--CGCCACCGCTGCCACGCTAA	629
OY	822	GAGAAATTTGGTGTGGTGGGTTGAGATTAACCTGSAACCCGAGCAGATCCGAGATCAAGTAT	881
Db	630	AGCAGGGTGTGCGCGCTGGGTGGTGCACCACTTCCAGAGACGGCAGCTGCTGCCAGGCTG	689
OY	882	TACAAATGAGCTAAAGCAAAAGCATTTGAATCCATACCTACGGGCAAGCAATTTGGAGAAAT	941
Db	690	TACCGGACAGCTGGCGCCTCATCTGCGGGTGAATGTCTTTGGCCGTTGCCAATGAGACGCCA	749
OY	942	GTCAAATATATAAAATTTGATTCCTACCAATATCGCTGTGAATTTTATCTTCTTTGAA	1001
Db	750	CTGTGCGCCAGCTGCTGCGTGGGCCACCGGGCCAGTACCGCTTACCTGTGCTTTTGAG	809
OY	1002	AATTCATATCCACAAGATTTTCATCATCAGCGAAAGACTAT--ACAAATGCTTTTGTGCTGGC	1058
Db	810	AACCTGTAGACACCGGACCTCATTTAGAGAAATTTGTGGCGCAACGCACTGGTGGCTGGC	869
OY	1059	TCGTACCTTGTCTTCTGGGACCATCTTAGGGAAAACATAAGAAATTAATTTCCAGCAAGT	1118
Db	870	ACTGTGCAATGTGGTGGGGGGCCCCCAGCGGCCACCTATGAGAGCCCTTCGTCGGGCTGAC	929
OY	1119	TCATTCATTCATGGAAGATTATTAATCTCCACAGTACGACTGAC	1162
Db	930	GCCCTTCTGATGTGATGACTTTGGCTCAGCCCGAGAGCTGGC	973
RESULT 9			
AAA35014			
ID	AAA35014	standard; DNA; 1701 BP.	
AC	AAA35014;		
XX			
DT	28-JUL-2000	(first entry)	
XX			
DE		Human adenosine receptor related polynucleotide seq ID NO:2703.	
XX			
KW		Human; adenosine receptor; low adenosine antisense oligonucleotide;	
KW		phosphotriphosphate; impaired respiration; inflammation; allergy;	
KW		allergic disease; bronchococcal infection; inhibitor; antiinflammatory;	
KW		anti-allergic; antihistaminic; cytoskeletal; analgesic; impaired airway;	
KW		lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;	
KW		respiratory distress syndrome; pain; cystic fibrosis; emphysema;	
KW		pulmonary hypertension; chronic obstructive pulmonary disease; COPD;	
KW		cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.	
XX			
OS	Homo sapiens.		
XX			
PN	MO200009525-A2.		
XX			
PD	24-FEB-2000.		
XX			
PF	03-AUG-1999;	99MO-DS17112.	
XX			
PR	03-AUG-1998;	98US-0095212.	
XX			
PA	(UYEC-) UNIV EAST CAROLINA.		
XX			
PI	Nyce JW;		
XX			

[illegible]

in vitro reactions to modify cell surface oligosaccharide(s) e.g. blood gp. determinants, to protect against transplant rejection

Claim 3; Fig 7; 249pp; English.

The sequence is that encoding human GDP-Fuc: [beta-D-Gal (1-4)]-D-GlcNAc alpha (1-3)-fucosyltransferase. The enzyme produced by the CC DNA may be non glycosylated. This prevents premature loss of enzyme activity. It can also be used in in vitro reactions to modify cell surface oligosaccharide moles. e.g. blood group determinants. See also AAO56906-12.

Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match 8.4%; Score 113.4; DB 15; Length 1654;

Best Local Similarity 51.2%; Pred. No. 3.3e-21;

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

```

OY 462 ATTCTGCTGTGGTGGGCGCATTTGGGCGAGACCTTTGACCTTACCTCTGCGCAAGCATG 521
DB 318 ATCTGCTGTGGAGCTGGCTTTTAAACAACCATAGCTCTGCCCCGCTGCTCAAGATG 377
OY 522 TT---CAACATCCAGATGCCATCTCACACGAGCGTTACCTGATACACAAATCCAT 578
DB 378 GTGCCGTGGACGGTGTACTCAACATCACTGCCGACCGAAGTGTATCCACAGGACAGC 437
OY 579 GCAGTCTGTATCCATCACCAGACATCAAGTTGGATCTGA--CAAAATTACCTCAGCA 635
DB 438 GCGGTATCTGTGCTACACCGAGAGGTGATGACACCCAGTGCCTCCAGCTCC 497
OY 636 GCTAGGCCACCTTCAGAAATGATTTGATGATGATGATGATGATGATGATGATGATGATG 695
DB 498 CCGAGGCGGACGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 557
OY 696 CAAAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 755
DB 558 CAGCTGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 617
OY 756 ATCCAAAGTGGCTTTAGGCT-----TCTGAGGATGATGATGATGATGATGATGATGAT 806
DB 618 ATCTTACAGCCCTTACGGCTGGCTGAGGCGGTGCTCCGCGACCGCTCCAGCTCC 677
OY 807 GAAAGTCCCAAGCAAGAAATTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
DB 678 AACCTCTGGCCAGACGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
OY 867 GCCAGATCACTATTACATGAGTAAAGCAAAAGCATTTGAATCCATTAACCTGAGCGCAA 926
DB 738 GCCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794
OY 927 GCATTGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 986
DB 795 CCGCTCCACAAAGCCCTGCGCCAGGAGAACATGATGAGAGCGCTCCCGGTACAAAGTTC 854
OY 987 TATCTTCTCTTGAATAATTCATCCACAAGATTTACATCAGCAAGAAACCTAT---ACAT 1043
DB 855 TATCTGCGCTTGGAGAACCTTTCACACCCGACATCATCAGCAAGAACTGTGGAGAAC 914
OY 1044 GCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
DB 915 GCCGTGAGAGCGCTGCGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
OY 1104 TATATTCACGATTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1163
DB 975 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
OY 1164 AAGTATCTGAGAGATGAGCAAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 1223
DB 1035 CGGTACTCTGAGAGCTGAGCAAGAGCAAGCGCCGCTACTGAGTACTTTCGCTGCGCG 1094
OY 1224 AAG 1226
DB 1095 GAG 1097

```

RESULT 13

AAT61680 standard; DNA; 1654 BP.

AAT61680;

19-JUN-1997 (first entry)

Human alpha(1,3)-fucosyltransferase (Fuc-TVI) DNA.

Alpha(1,3/1,4)-fucosyltransferase; Lewis enzyme; Fuc-TVI; glycosylation; oligosaccharide; blood group; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 129..1208

W09709421-A1.

13-MAR-1997.

06-SEP-1996; 96MO-US13816.

08-SEP-1995; 95US-0525058.

(UNMI) UNIV MICHIGAN.

Legault DJ, Lowe JB;

WPI; 1997-192897/17.

P-PSDB; AAW13643.

New recombinant fucosyltransferase proteins - useful for modifying cell surface oligosaccharide structures

Example 6; Page 291-292; 329pp; English.

A DNA clone (AAT61680) codes for an enzyme (AAW13643) that functions as a GDP-Fuc:beta-D-gal(1,4)-D-GlcNAc alpha(1,3)-fucosyltransferase (Fuc-TVI) which can utilize neutral type II oligosaccharide precursors but not type I glycoconjugates. It was identified in a human lambda phage genomic DNA library using a probe obd. from Lewis enzyme cDNA (see also AAT61675). The Fuc-TVI DNA can be used to construct animal cell lines with specific capabilities with respect to post-translational modification of the oligosaccharides of expressed proteins or lipids, to produce recombinant Fuc-TVI for use in oligosaccharide prodn., to genotype individuals at this fucosyltransferase locus, and in the prodn. of chimeric fucosyltransferases (see also AAW14512-32).

Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match 8.4%; Score 113.4; DB 18; Length 1654;

Best Local Similarity 51.2%; Pred. No. 3.3e-21;

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

```

OY 462 ATTCTGCTGTGGTGGGCGCATTTGGGCGAGACCTTTGACCTTACCTCTGCGCAAGCATG 521
DB 318 ATCTGCTGTGGAGCTGGCTTTTAAACAACCATAGCTCTGCCCCGCTGCTCAAGATG 377
OY 522 TT---CAACATCCAGATGCCATCTCACACGAGCGTTACCTGATACACAAATCCAT 578
DB 378 GTGCCGTGGACGGTGTACTCAACATCACTGCCGACCGAAGTGTATCCACAGGACAGC 437
OY 579 GCAGTCTGTATCCATCACCAGACATCAAGTTGGATCTGA---CAAATTACCTCAGCA 635
DB 438 GCGGTATCTGTGCTACACCGAGAGGTGATGATGATGATGATGATGATGATGATGATGATG 497
OY 636 GCTAGGCCACCTTCCAGAAATGATTTGATGATGATGATGATGATGATGATGATGATGATG 695

```

```

Db      498 CCGAGGGGGAGGGGGGCGCATGGATCTGGTTCAGCATGGAGTCCCAAGCCACTGCTGG
QY      696 CAAAAGAGTGGCATTTGAGCACTGTTTAACTGACTCTGACTTACCCCGTGATTCAGAT
Db      558 CAGCTGAAACCATGAGCGGATCTTCAATCTCAACCATGTCCTACCGCAGAGACTCGGAC
QY      756 ATCCAACTGCTTATGGCT-----TCTTGACGTTAAACACAAATCCCTTGCTTT
Db      618 ATCTTCACGGCTTACGGCTGAGCCGCTGGTCCGCGCCAGCTTCCCAACCCACCTCTC
QY      807 GAAGTGGCAAGCAAGAAGAAATTTGTTGGTGGTGTGAGTAACTGGAAACCCGAGCAT
Db      678 AACCTCTGGCCAAAGACCGAGCTGTGTGGCTGGGCACTGTCCAGCTGGGGCCAACTCC
QY      867 GCCAGACTCAAGATTTACAAATGAGCTAAGCAAAAGCATTTGAAATCCATCTACGAGCA
Db      738 GCCAGGGTGGCTACATCAAGAGAGCTGACAGCCCATCTCAAGGGGAGACGTACGG---A
QY      927 GCATTTGGAATATGTCATGATTAATAATTTGATTCCTTCAATATCTGCTTGAATTT
Db      795 CGCTCCCAAGAGCCCTGCCCCAGGAAACCATGATGAGAGCGCTGCCGTACAAAGTTC
QY      987 TATCTTCTTTGAAATTTCAATCCACAGAGATTACATCCAGGAAAGCTAT---ACAAAT
Db      855 TATCTGGCTTCGAGAACTCTTGACCCCGACTACATCAACGGAAGCTGTGAGAGAC
QY      1044 GCTTTTCTGGCTGCTGCTGACTGTGTTCTGGAGCACTAGAGCAAACTATGAGAT
Db      915 GCCCTGAGGCTGTGGGCGGTGCGGTGTGTGTGGCCCGCCAGCAAGCAACTAGAGAG
QY      1104 TATATTCACAGATTCATTCATTCATGTGAGAAATTAATTAATCTCCAGTACAGCA
Db      975 TTCTCTGACACCCAGCGCTTATCACTGAGAGACTTCCAGAGCCCAAGAGACTGGCC
QY      1164 AAGTATCGAAGAGAGCGCAAAACCAATAAGTTAATCTAGTTAGTTAATCTAGAG
Db      1035 CGGTACTGCAAGAGCTGGAAGAGACCAAGCCCGCTACTGAGTACTTTCGTGGCG
QY      1224 AAG 1226
Db      1095 GAG 1097

RESULT 14
AAQ56912
ID      AAQ56912 standard; DNA; 1086 BP.
XX
AC      AAQ56912:
XX
XX      26-JUL-1994 (first entry)
DE
XX      Lewis blood group fucosyltransferase DNA.
XX
KM      Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
XX      surface; oligosaccharide; ss.
XX
XX      WO9402616-A.
XX
XX      03-FEB-1994.
PD
XX      20-JUL-1993; 93MO-US06703.
PF
XX      20-JUL-1992; 92US-0914281.
PR
XX      (UNMI ) UNIV MICHIGAN.
PA
XX      Lowe JB;
PI
XX      WPI; 1994-04874/06.
XX
XX      DNA fragment encoding a glycosyltransferase - can be used for in
PT      vitro reactions to modify cell surface oligosaccharide(s) e.g.

```

```

PT      blood gp. determinants, to protect against transplant rejection
XX
PS      disclosure; Fig 6; 249pp; English.
XX
CC      The sequence is that encoding a Lewis blood group fucosyl
CC      transferase (Fuc-III). The enzyme produced by the DNA may be
CC      non glycosylated. This prevents premature loss of enzyme activity.
CC      It can also be used in vitro reactions to modify cell surface
CC      oligosaccharide moles. e.g. blood group determinants.
CC      See also AAQ56905-11.
XX
SQ      Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;
Query Match      8.3%; Score 112.2; DB 15; Length 1086;
Best Local Similarity 51.6%; Pred. No. 5,7e-21;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY      537 TGGCATCTCAGAAAGGACCTTCACTGATACAAATAATCCATGCACTTGTGATCATCAG
Db      271 TGGCAATCACTGCGCCGCAAGAGTGTACCAAGGACAGACAGGTCATCTGACACAC
QY      597 CGAGACATCAGTTGGGA---TCTGACAAATTTACCTGAGCAAGCTGAGCCACCTTCCAG
Db      331 TGGGATATCTATGTCCAAACCTTAAGTCAAGCTTCCACCTTCCCGAGGCGCAGGGAG
QY      654 AATGATTTGATGATTTGGAATTCACCACTCACACTCCCAAAAGAGTGGCATTTAG
Db      391 CGCTGATCTGTGTTCAACTTGGAGCCACCCCTTAATCTCCAGACACCTGAGAC
QY      714 CACTTGTTAACCTGACTGACTTACCGCGGTGATTCAGATATCCAAATGCTTATGGC
Db      451 AGATTAATCTCAATCTCAACATGCTCTTACCGAGACTCCGACATCTTCAAGCCCTTAC
QY      774 TTCTTGACGGTAAAGCAAAATCCCTTGGT-----GTTTAAAGGCCAAACAAAGAG
Db      511 TGGCTGAGCGCGTGTGCGGCCAGCTGCCACCCAGCCGCTCAACCTTGTGGCCAAAGC
QY      825 AATTTGTTGTGTTGGTTGAGTAATCTGAAACCTTGAGCAGTCCAGAGTCAAGTATAC
Db      571 GAGCTGTGCTGTGGGCGGTGTCCAACTGGAACCGGACATCAACGAGGTGGCTTACTAC
QY      885 AATGAGCTAAGCAAAAGCATTTGAATTCATCTAGAGGCGAGACATTTGGAAATATGTC
Db      631 CAGAGCCTGAGGCTCATCTCAAGGTGAGGTACGAGAGCTCCCAAGACCCCTGCGCC
QY      945 AATGATTAATAATTTGATCTTACCATATCTGCTGTAATTTTATCTTCTTGAATAAT
Db      691 AAGGGAGCCA---TGATGGAGAGCTGTCTCCGCTACAGTTTCTACTTGGCTTGGAGAC
QY      1005 TCAATCCACAAAGATTTACATCAGGAAAGCTAT---ACAAATGCTTTTGTGGCTGCTC
Db      748 TCCTTGACACCCGACATCAATCAGGAAAGCTGTGAGAGAACGCCCTGAGGCTGGGCC
QY      1062 GTACCTGTTTCTTGGGACCATCTAGGAAAACCTATGAGAAATTAATTCAGCAGATTCA
Db      808 GTGCCCGGTGTGCTGGGCCCAAGCAAGCAATAGAGGTGTCTGTGCCACCCGAGCGC
QY      1122 TTCAATCATGTGGAAGATTTAATCTCCAGAGTGTAGCAAGTATCTGAGAGAGTC
Db      868 TTCAATCAGCTGAGACATCTCCAGAGCCCAAGAGCTGTGCGGTACTGCGAGAGCTG
QY      1182 GACAAAAACAAATAAGTATTAATCTTAACTTAACTGAGAGAG 1226
Db      928 GACAAAGAGACAGCCCGCTACCTGAGGTACTTTCGCTGGCGGAG 972

RESULT 15
AAT80111
ID      AAT80111 standard; cDNA; 2042 BP.
XX
AC      AAT80111:
XX
XX      01-DEC-1997 (first entry)

```

XX Fucosyltransferase III coding sequence.

DE P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;

XX calcium-dependent carbohydrate binding protein; platelet; endothelium;

KW thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;

KW inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;

KW disseminated intravascular coagulation; rheumatoid arthritis; antibody;

KW adult respiratory distress syndrome; tumour metastasis;

KW fucosyltransferase III; acetylglucosaminyltransferase;

KW autoimmune disorder; inflammatory disorder; ss.

XX Homo sapiens.

OS MO9706176-AZ.

PN

XX 20-FEB-1997.

PD

XX 02-AUG-1996; 96MO-US12820.

PF

XX 17-MAY-1996; 96US-0649802.

PR 03-AUG-1995; 95US-0510920.

PR 15-MAY-1996; 96US-0017794.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Cummings RD, Mcever RP, Moore KL;

PI WPI; 1997-154206/14.

XX

XX New O-glycan(s) derived from P-selectin glyco:protein ligand

PT (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.

PT for treating inflammatory or autoimmune disorders or tumours

PT

XX Disclosure; Page 80-81; 96pp; English.

XX

XX This sequence represents the coding sequence for the human

XX fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-

CC acetylglucosaminyltransferase encoded by AATG0112 are needed for the

CC expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in

CC mammalian cells. P-selectin is a calcium-dependent carbohydrate binding

CC protein expressed on the surfaces of activated platelets and endothelium

CC in response to thrombin and other agonists. PSGL-1 (see AAW26174) is a

CC high affinity P-selectin ligand produced by leukocytes. Binding of

CC P-selectin to PSGL-1 is calcium ion dependent and is abolished by

CC treatment of the ligand with sialidase. PSGL-1 is a homodimer, with a

CC highly extended extracellular domain, which is a feature of mucin-like

CC proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of

CC the invention (see AAW26175-W26188). PSGL-1 is also targeted by the

CC O-glycans, and O-glycanated products of the invention. The products are

CC used for inhibiting the binding of P-selectin and other selectins to

CC PSGL-1. They can be used for the treatment of inflammation, injury

CC resulting from ischaemia and reperfusion, bacterial sepsis and

CC disseminated intravascular coagulation, adult respiratory distress

CC syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and

CC other autoimmune or inflammatory disorders. The products can also be used

CC in the production of antibodies and in diagnostic applications.

XX

XX

XX Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

Query Match 8.3%; Score 112.2; DB 18; Length 2042;

Best Local Similarity 51.6%; Pred. No. 8e-21;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY 537 TGGCATCTCCACAGCAGCTTACGTACAAACAAATCCCATGTCATGATCATAC 596

DB 343 TGGCAGATCTCTCCGACGACGAGTGTACCCACAGGACACAGGTCATGTGACACAC 402

OY 597 CGAGACATCAAGTTGGGA--TCTGACAAATTTACTCTAGCAGACTAGCCCTTCAG 653

DB 403 TGGGATCTCTATGTCCAAACCTTAAGTCAAGCTTCCCTCCGAGGCGCAGGGGCGAG 462

OY 654 AATGATTTTGAATTTGAATGACCAACTCACACTCCCAAAAAGAGTGGCATTTGAG 713

DB 463 CCGTGGATGTGTTCAACTTGGAGCCACCCCTTAACCTGCGACACCTGTAACCCCTGAC 522

OY 714 CACTGTTTAACTGACCTGACCTGACCTGACCTGATTCAGATATCCAAAGTGCCTTATGAC 773

DB 523 AGATACCTTAAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 582

OY 774 TTTCTGACCTTAAGCACAATCCCTTCGT-----GTTGAAGTGCAGCAAAAGAG 824

DB 583 TGGCTGACCCGTGGGCGGCGGACAGCTGCGCCACCCAGCCGCTCAACCTCTCGGCGCAAGAC 642

OY 825 AATGCTGTGCTGGGTTGTGTGAGTAACTGAGCAACCTGAGCATGCGAGTCAAGTAAATAC 884

DB 643 GAGCTGTGCTGGGCGGCGGCTGTCAACTGGAAGCGGAGCTCAAGGAGGTGCGCTACTAC 702

OY 885 AATGAGCTTAAGCAAAAGCACTTGAATCCATACCTACGAGGCAAGCATTTGGAAATATGTC 944

DB 703 CAGAGCTTCAGAGCTCATCTCAAGGTGACGTGTACGAGCGCTCCACAGCCCTGCGCC 762

OY 945 AATGATTAATAATTTGATTTCTTCCATATCTGCTTGAATTTTATCTTCTTTGAAT 1004

DB 763 AAGGGGACCA--TGATGAGAGAGCTGTCCGCTCAAGTTCTACCTTGCGCTTGAGAGAC 819

OY 1005 TCAATCCACAGGATTTACATCAAGGAAAGCTAT--ACAATGCTTTTCTGCGTGGCTCT 1061

DB 820 TCCCTGACCCCGGACATCACTCAAGAGAGCTGTGAGAGAACCCCTGAGGCGTGGGCGC 879

OY 1062 GTACCTGTGTTCTGGGACCATCTAGGGAAACATATGAAATTAATTCAGACAGATTCA 1121

DB 880 GTGCGCGTGTGCTGTGGGCCCGACAGAGCACTACAGAGAGTCTCCACACCGCAGCGC 939

OY 1122 TTCAATTCATGTGAAGATATATTAACCTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1181

DB 940 TTCAATTCATGTGAAGATATTAACCTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 999

OY 1182 GACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1226

DB 1000 GACAAACAAACAGCCCGGCTACCTGAGCTTCTTCCCTGCGGCGAG 1044

Search completed: October 5, 2002, 23:16:51

Job time: 182.049 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 1126.97 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: US-09-744-748-5_COPY_29_1374

Perfect score: 1346
Sequence: 1 actgcctccctgcagctc.....agaatggttttgaaattaa 1346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_estlum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vtc:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	573.8	42.6	697	9 BB397507	BB397507 BB397507
2	504.2	37.5	568	10 BM021317	BM021317 1e75f03.Y
3	492.8	36.6	607	10 B1102760	B1102760 602888338
4	422	31.4	716	9 AV725071	AV725071 AV725071
5	387.2	28.8	648	9 BB595164	BB595164 BB595164
6	375	27.9	625	9 BB649967	BB649967 BB649967
7	363.2	27.0	767	9 BB640911	BB640911 BB640911
8	324.8	24.1	649	9 AV338075	AV338075 AV338075
9	322.2	23.9	658	9 BB651409	BB651409 BB651409
10	307.4	22.8	579	10 BM021081	BM021081 1e75f03.x
11	296.8	22.1	1101	12 CNS0595T	AL327638 Tetraodon
12	239.2	17.8	670	9 BB626126	BB626126 BB626126
13	212.2	15.8	533	10 B1604066	B1604066 603243886
14	184.2	13.7	570	10 BM082484	BM082484 f125608.x
15	177	13.2	1100	12 CNS05792	AL324368 Tetraodon
16	170.2	12.6	1016	12 CNS0429P	AL271222 Tetraodon
17	169.4	12.6	335	12 A0054539	A0054539 CIT-HSP-2

18	169	12.6	617	10 BE266935	BE266935 601189956
C 19	165.2	12.3	487	12 A0246164	A0246164 HS-2060.A
C 20	165	12.3	844	12 CNS018V	AL168808 Tetraodon
C 21	157.4	11.7	887	12 CNS0200S	AL208693 Tetraodon
C 22	155.4	11.5	651	10 BJ036855	BJ036855 BJ036855
C 23	151.6	11.3	666	10 BJ050300	BJ050300 BJ050300
C 24	151.2	11.2	711	12 CNS034XF	AL228012 Tetraodon
C 25	151	11.2	501	12 B41073	AL1073 HS-1052-B2-
26	149.6	11.1	367	9 BB843869	BB843869 BB843869
27	146.6	10.9	640	9 BJ073227	BJ073227 BJ073227
28	136.2	10.1	199	9 BB398103	BB398103 BB398103
C 29	123.4	9.2	674	10 BE605503	BE605503 f110104.x
C 30	120.6	9.0	971	12 CNS047TV	AL278032 Tetraodon
C 31	120.4	8.9	120	10 B1326811	B1326811 STR00078
C 32	117.8	8.8	1023	12 CNS028ANG	AL188773 Tetraodon
C 33	114.8	8.5	661	9 BB651647	BB651647 BB651647
C 34	102.8	7.6	487	10 BE663396	BE663396 UT-M-BH-
C 35	100.2	7.4	702	10 BJ043578	BJ043578 BJ043578
C 36	93.8	7.0	863	12 CNS02ROF	AL210840 Tetraodon
C 37	90.4	6.7	720	10 BF180380	BF180380 601804564
C 38	87.4	6.7	551	10 BJ035701	BJ035701 BJ035701
C 39	87.4	6.5	601	10 BJ036679	BJ036679 BJ036679
C 40	87.4	6.5	605	10 BJ028804	BJ028804 BJ028804
C 41	87	6.5	1056	12 CNS01VL7	AL169232 Tetraodon
C 42	86.8	6.4	304	9 A0081478	A0081478 A0081478
C 43	84.6	6.3	102	12 B70664	B70664 CIT-HSP-206
C 44	84.2	6.3	638	10 BJ035697	BJ035697 BJ035697
45	83.6	6.2	1057	12 CNS03ADC	AL235065 Tetraodon

ALIGNMENTS

RESULT 1
BB397507
LOCUS
DEFINITION
BB397507 RIKEN full-length enriched, ES cells Mus musculus cDNA clone C330004C19 3' similar to AB015426 Mus musculus Fut9 mRNA for alpha1,3-fucosyltransferase IX, mRNA sequence.

ACCESSION
BB397507
VERSION
BB397507.2 GI:16411808

KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus

REFERENCE
A0081478
A0081478

1 (bases 1 to 697)
Mammalia; Eutheria; Rodentia; Sclurognathia; Muridae; Murinae; Mus.

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)

On Jul 15, 2000 this sequence version replaced gi:9216903.

Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

TEL: 81-45-503-9222
FAX: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10):1617-1630 (2000)

wag1, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura

Tue Oct 8 10:18:36 2002

us-09-744-748-5_copy_29_1374.rst

Page 4

Db 599 GAAACTAT 607

RESULT 4
AV725071
LOCUS AV725071
DEFINITION HRC Homo sapiens cDNA clone HRCBLB02 5', mRNA sequence.
ACCESSION AV725071
VERSION AV725071.1 GI:10830091
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.
Homo sapiens cDNA HRC clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

Source

1. 716
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HRCBLB02"
/clone_id="HRC"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); site_1: EcoRI; site_2: XhoI"
BASE COUNT 219 a 113 c 133 g 249 t 2 others
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 422; DB 9; Length 716;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 925 AAGCATTTGGAGATATGCAATGATATTAATAATTTGATCTTACCATTATGCTTGAAT 984
DB 1 AAGCATTTGGAGATATGCAATGATATTAATAATTTGATCTTACCATTATGCTTGAAT 60
QY 985 TTTATCTCTCTTGAATAATTCATCCACAGAGATTACATCAGGAAAGCTTATACATG 1044
DB 61 TTTATCTCTCTTGAATAATTCATCCACAGAGATTACATCAGGAAAGCTTATACATG 120
QY 1045 CTTTTCGGCTGCTGTGACTGTGCTGCTGGACCATCTAGGAAACTATAGAGAAAT 1104
DB 121 CTTTTCGGCTGCTGTGACTGTGCTGCTGGACCATCTAGGAAACTATAGAGAAAT 180
QY 1105 ATATTCAGCAGATTCATTCATTCATGATGAGGATTATATCTCCACAGGACTAGCA 1164
DB 181 ATATTCAGCAGATTCATTCATTCATGATGAGGATTATATCTCCACAGGACTAGCA 240
QY 1165 AGTATCGAAGAGATGACAAAACAAATTAATTAATCTTACTTAACTGAGAGA 1224
DB 241 AGTATCGAAGAGATGACAAAACAAATTAATTAATCTTACTTAACTGAGAGA 300
QY 1225 AGGATTCAGTAAATCTTCACAGATTTGGGAATCACAATGATGCTTGGCTGGATC 1284
DB 301 AGGATTCAGTAAATCTTCACAGATTTGGGAATCACAATGATGCTTGGCTGGATC 360
QY 1285 ATGTGAAAAGCATCAAGATATTAAGTCTGTGTAATTTAGAAATGCTTTGGAAT 1344
|||||

Db 361 ATGTGAAAAGCATCAAGATATTAAGTCTGTGTAATTTAGAAATGCTTTGGAAT 420

QY 1345 AA 1346
DB 421 AA 422

RESULT 5
BB595164
LOCUS BB595164
DEFINITION BB595164 RIKEN full-length enriched, adult male corpus striatum Mus musculus cDNA clone CO30004N01 5', mRNA sequence.
ACCESSION BB595164
VERSION BB595164.2 GI:16450167
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Koda, M., Koye, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, K., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, T., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
On Nov 30, 2000 this sequence version replaced gi:11491766.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wegli, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanka, I., Alawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Location/Qualifiers
1. 648
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

			/clone="C030004N01"	/clone.lib="RIKEN full-length enriched, adult male corpus striatum"
			/sex="male"	/tissue_type="corpus striatum"
			/dev_stage="adult"	/lab_host="DH10B"
			/note="Site.1: Sali; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15'	
			GAGGAGAGAAGTCCACAGCGCTTTTTTTTTTTTVVN 3'). cDNA was prepared by using trehalose thermo activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGATTCGAGGTATTAAATTAATCCCCCCCCCCC 3'). cDNA was cloned into the xhoI and BamHI sites. Vector: a modified Bluescript KS(+) after bulk excision from Lambda PUC I. Cloning sites, 5' end: Sali; 3' end: BamBI"	
BASE COUNT	141 a	206 c	138 g	163 t
ORIGIN				
Query Match	28.8%	Score 387.2;	DB 9;	Length 648:
Best Local Similarity	80.4%;	Pred. NO.1.7e-86;		
Matches 479;	Conservative	0;	Mismatches 113;	Indels 4;
				Gaps 2.
QY	10	CCCTGCAGCTCCCGCGGCCCGCGCGTGTCCTCTCGGTGTCGCCAGCCCGCAGTCG	69	
Db	57	CCCGGTGCTTGAAGACC CGCGCGCGCTGTGGCTCTTGGCTACCCCGCGAGATCCGGCTA	116	
QY	70	CGCTTTAGGACAGCGCCGCCACCGCGCTGCGCCCTGCTGCTCGCGCGGCAAC	129	
Db	117	CGTACTGTGAGCAGCGCGCCGACACCGCACCGCACCGCTCG--TCCCGCGCGCGCGGCTGC	174	
QY	130	CCTGCGAGCGCCCGGAGTGGCGCTTACCCCTTAGCGAGATTTGAAATGTATAATCA	189	
Db	175	CTCTCCCGCGCCCTGGATGGAGCTTTCCTCGCTGGGATGCTTTGAAATGTATAATCA	234	
QY	190	AGGATTTGATATAATACAGTAGAAGTAGATATAACAACGTGTACGTGCTTCCCATGATATGT	249	
Db	235	CGCATCTGATAGACCGCGAGTAGTGTACAGCACTCATCT--CTGCTTTCATGATATGT	292	
QY	250	CTCTAATATGAAAAATTTATGACATCAACATCCAAGAAATTTTGGCCCATTTTAATG	309	
Db	293	CTTCACTGAAAAATTTATGACATCAACATCCAAGCAATTTTGGCCCATTTTAATG	352	
QY	310	TCTGCAATATCCGTGGCTGTCTCATGGAGTCTCTCATTTACATCAAACCTTACACA	369	
Db	353	TCTGCAATATCCGTGGCTGTCTCATGGAGTCTCTCATTTATATCAAGCCACACACA	412	
QY	370	GCTGATATTAGTCAATGGAATGACGACGCTGTGCTGAAAATGAAAAATCTCTTT	429	
Db	413	GCTGATATTAGTCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT	472	
QY	430	CCACCAAAACGATATTTTATGAAACATACATCTCGGTGGGTGGGTGGCCATTTGGGC	489	
Db	473	CCACCAAAACGATATTTTATGAAACATACATCTCGGTGGGTGGGTGGGTGGGTGGGC	532	
QY	490	AGACCTTTGACCTTACATCTGCGCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAA	549	
Db	533	AGACCTTTGACCTTACATCTGCGCAAGCAATGTTCAATATCCAGGAGTGCATCTCACAA	592	
QY	550	CGGACCGTCTACGTACAAACAAATCCCATGCAATGTCATCATCAACGAGACATC	605	
Db	593	CAGACCGTCTCTGACAAACAAATCCCATGCGGTCTGATGATACCATATATGATC	648	

	RESULT	6			
	BBS49967				
LOCUS	BBS49967				
DEFINITION	BBS49967 RIKEN full-length enriched, 16 days embryo head Mus musculus CDNA clone C13009N17.5, mRNA sequence.	625 bp	mRNA	linear	EST 26-OCT-2001
ACCESSION	BBS49967				
VERSION	BBS49967.1				
KEYWORDS	EST,				
SOURCE	Mus mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I (bases 1 to 625)				
AUTHORS	Arakawa,T., Carrincci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himemoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Koude,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasak,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota.T., Muramatsu,M., and Hayashizaki.Y.				
TITLE	RIKEN Mouse ESTS (Arakawa,T., et al. 2001)				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Yoshihide Hayashizaki				

FEATURES
source
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.
Location/Qualifiers
1. .625

```
/sex="mixed"  
/tissue.type="head"  
/dev_stage="16 days embryo"  
/lab_host="DH10B"  
/note="Site:1: Salt; Site:2: BamHI. cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken
```


Db	510	CCACAAACATGATATATGTAAACGAACATACATCTGCTGTTAGTATGACCATWTGGCC	569
Qy	490	AGACCTTGACCTTACATCTCCGACGAAGTTTAAACATCCAAAG	535
Db	570	AGACCTTGACCTTACATCTCCGACGAAGTTTAAACATCCAAAG	615
RESULT 10			
LOCUS	BM021081	579 bp	mRNA
DEFINITION	1e75f03.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens		
VERSION	cDNA 3, similar to TR:09Y231 Q9Y231 ALPHA-3-FCOOSYLTRANSFERASE. ;,		
KEYWORDS	mRNA sequence.		
ACCESSION	BM021081		
VERSION	BM021081.1	GI:16535437	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 579) Melton,D., Brown,J., Kenty,G., Permitt,A., Lee,C., Keastner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wyle,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Konko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y., and Howers,Y.		
TITLE	Endocrine Pancreas Consortium		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Douglas Melton, Klaus H. Keastner, & Hiroshi Inoue Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@ionp.harvard.edu Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) High quality sequence stop: 436.		
FEATURES	Location/Qualifiers		
SOURCE	1..579		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"		
	/sex="Both"		
	/tissue_type="Islets of Langerhans"		
	/dev_stage="Adult"		
	/lab_host="DH10B"		
	/note="Organ: Pancreas; Vector: pSPORT1; Site.1: Not 1; Site.2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 genome Research 6:791-806; 0.5 microgram single-stranded library Plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."		
BASE COUNT	203 a	103 c	93 g
ORIGIN	180 t		
Query Match	22.8%;	Score 307.4;	DB 10; Length 579;
Best Local Similarity	99.7%;	Pred. No. 1.9e-66;	
Matches 308; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
1038	TACAAATGCTTTTGGCTGGCTCTGTACCTGTTGTTCTGGAGCATTAGGAAACTAT 1097		

	Db	579	TACAAATGCTTTTCGTGGCGCTCTCTACTTGTTGTTGGGCCAATCTAGGGGAANAACATAT	520
OY	1098	GAGAATTATATTCGCAGCAGATTCATTCAATCATGTCGAAGAATTATTAACCTCCAGTAG	1157	
Db	519	GAGATTATATTCGACGACGAGATTCAATTCATGTCGGAAGAATTATAACTCTCCAGTAG	460	
OY	1158	CTAGCAAAAGTATCTGAGGAAGTGCACAAAAACAATTAAGTTAACCTTACTTTAAC	1217	
Db	459	CTAGCAAAAGTATCTGAGGAAGTGCACAAAAACAATTAAGTTAACCTTACTTTAAC	400	
OY	1218	TGGAGAGAGATTCACCTGTAAATCTTCACAGATTTTGGGAATCACATGCTTTGGCT	1277	
Db	399	TGGAGAGAGATTCACCTGTAAATCTTCACAGATTTTGGGAATCACATGCTTTGGCT	340	
OY	1278	TGCGATCATGTGAAAAGCATCAACAATATTAAGTCTGTGTAATTTAGCAAATGCTTT	1337	
Db	339	TGCGATCATGTGAAAAGCATCAACAATATTAAGTCTGTGTAATTTAGCAAATGCTTT	280	
OY	1338	TGGAATTA 1346		
Db	279	TGGAATTA 271		
RESULT 11	CNS059ST/c	1101 bp DNA linear GSS 26-MAY-2000		
LOCUS	CNS059ST/c			
DEFINITION	Tetradon nigroviridis genome survey sequence T3 end of clone 04P22 of library A from Tetradon nigroviridis, genomic survey sequence.			
ACCESSION	ALJ27638			
VERSION	ALJ27638.1 GI:8221227			
KEYWORDS	GSS; genome survey sequence.			
SOURCE	Tetradon nigroviridis.			
ORGANISM	Tetradon nigroviridis. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neoceratrygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphae; Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodon.			
REFERENCE	1 (bases 1 to 1101) Roest-Crollius H., Jallion O., Dasilva C., Fitzames C., Fisher C., Bouneau L., Billault A., Quetier F., Saurin W., Bernot A. and Weissenbach J.			
AUTHORS	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis Unpublished			
TITLE	2 (bases 1 to 1101)			
JOURNAL	Roeft-Crollius H., Jallion O., Dasilva C., Bouneau L., Fisher C., Bernot A., Fitzames C., Wincker P., Brottier P., Quetier F., Saurin W. and Weissenbach J.			
REFERENCE	Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence Unpublished			
AUTHORS	3 (bases 1 to 1101)			
JOURNAL	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases			
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .			
FEATURES	location/Qualifiers			
source	1..1101			
	/organism="Tetradon nigroviridis"			
	/db_xref="taxon:99883"			
	/clone="04SP22"			
	/clone_id="A"			
	/note="Genoscope sequence ID : COAA045DH11Al-end : T3"			
BASE COUNT	287 a 249 c 269 g 278 t 18 others			
ORIGIN				
Query Match	22.1%	Score 296.8;	DB 12;	length 1101;
Best Local Similarity	58.3%	Pred. No. 1.le-63;		


```
OY 522 TTCAACATCCAAAGGATGCATCTCACAACGACCGGTCACTGTACACAAATCCCATGCA 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 713 TACAAATATGACGGCTGTGAGCTCAGATGATGCCCTTACGACAAAGCGGAGGCT 654
OY 582 GTTCTGATCCATCCACGAGACATCATGTTGGATGTGACAAATTTACCTCAGCAAGCTAGG 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 653 GTTCTGTCTTCCCAAGAGACATACAGTGGAAATCTGGAAACTTCCGAGTGAGCCGCGG 594
OY 642 CCACCCCTCCGAAATGGATTTGGATTTGGAAATCCACCACTCACACTCCCAAAAG 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 CCATATTTTCAGAGGTGATTTGGTTCTACCTGGAATCTCCAGAGAACACAAATCAGGATA 534
OY 702 AGTGGCAATGAGCACTGTGTTAACTGACTGTGACTTACCGCGGTGATTCAGATATCCAA 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 CCAGGCCCTGGAACAGTTTCCATGATCTGTAACACAGAAAGATTCGACATTTGTA 474
OY 762 GTGCTTA---TGCTCTTTCAGCGTAAAGCACAATCCCTCGTGTGTAAGTCCCAAGC 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 GCCAGGTACCTCTTACCATCAGAGAGAAATACTGACGGAGAAATATGTTCTGCCGAG 414
OY 819 AAAGAGAAATTTGTTGCTGGTTGTGATTAAGTGAACCCCTGAGC-----ATGCCAGA 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 AAAAACAACCTGGCTGTGGATGTAGCAATGCTACATCCACTGGGACAGTACAAGA 354
OY 873 GTCAAGTATTCATGAGCTAAGCAAAAGCATTTGAATCCATACCTACGGCAAGCATTT 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 GCGCAATTTTATTAATGAACATATCCAAACATATTAAGTGTGTTGGTGTGTGAC 294
OY 933 GGAGATATATGTCATTAATTAATAAATTTGAT---CCTACCATATCTGGTTGTAATTTTAT 989
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 ACAGGGGTGAGGTGGAGATTGATCAGTATTTACCGACCGTTGCTAGCTGTAGTTTAC 234
OY 990 CTTTCCTTTGAAATTCATTCACAGGATTCATCAAGGAAAGCTATACATGCTTTT 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 TTATCATTTTGAAGACTCATTTACACAAGATTACATCAAGAGAGTCAACGNAACCCCT 174
OY 1050 CTGGCTGGCTGTACCTGTGTCTGGGACCATCTAGGGAAACATAGAGATTATATT 1109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 GCACGAGGAACCGTTCCTGTAGTTCCTGGACCCCAAGAGCCAACCTAGCAATGTCTTC 114
OY 1110 CCACGAGATTCATTCATGTGAGAGATTTAATCTCTCCAGTGAAGCTAGCAAGTAT 1169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 CCTTCAGACTCTTATATTCAGTATGATTTTTCAGATCCAAAGGCACTGGCAGAGCAT 54
OY 1170 CTGAAGGAGTGCAGAA-AAACAATAAGTATATACCTTAGTTACTTTAACTGGA 1221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 CTGCTCTCCTTGAGACAAGNATGAGAGGCATACATGGCTACTTGTAGTGA 1
```

Search completed: October 6, 2002, 07:16:35
Job time : 1134.97 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 5, 2002, 22:42:00 ; Search time 27.4635 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-5_COPY_29_1374

Perfect score: 1346
Sequence: 1 actgtctccctcgcagctc.....agaaatgttttgaaatcaa 1346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 38353 segs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	127	9.4	1814	2	US-08-483-151-1
2	127	9.4	1814	2	PCT-US96-06427-1
3	113.4	8.4	1654	1	US-07-914-281-13
4	113.4	8.4	1654	1	US-08-393-246-13
5	113.4	8.4	1654	1	US-08-525-058A-13
6	113.4	8.4	1654	2	US-08-696-731-13
7	113.4	8.4	1654	4	US-09-042-531-13
8	112.2	8.3	1086	1	US-07-914-281-12
9	112.2	8.3	1086	1	US-08-393-246-12
10	112.2	8.3	1086	1	US-08-525-058A-12
11	112.2	8.3	1086	2	US-08-696-731-12
12	112.2	8.3	1086	2	US-09-042-531-12
13	112.2	8.3	2043	3	US-09-063-237-2
14	112.2	8.3	2043	1	US-07-914-281-1
15	112.2	8.3	2043	1	US-08-393-246-1
16	112.2	8.3	2043	1	US-08-525-058A-1
17	112.2	8.3	2043	1	US-08-696-731-1
18	112.2	8.3	2043	2	US-08-696-731-1
19	112.2	8.3	2043	4	US-09-042-531-1
20	112.2	8.3	2043	5	PCT-US91-00899-1
21	111.8	8.3	1316	1	US-07-914-281-10
22	111.8	8.3	1316	1	US-08-393-246-10
23	111.8	8.3	1316	1	US-08-525-058A-10
24	111.8	8.3	1316	2	US-08-696-731-10
25	111.8	8.3	1316	4	US-09-042-531-10
26	100.8	7.5	1215	5	PCT-US91-00899-13
27	100.8	7.5	1215	5	PCT-US91-00899-5

28	100.8	7.5	1488	1	US-07-914-281-9	Sequence 9, Appl1
29	100.8	7.5	1488	1	US-08-393-246-9	Sequence 9, Appl1
30	100.8	7.5	1488	1	US-08-525-058A-9	Sequence 9, Appl1
31	100.8	7.5	1488	2	US-08-696-731-9	Sequence 9, Appl1
32	100.8	7.5	1488	4	US-09-042-531-9	Sequence 9, Appl1
33	100.8	7.5	2175	4	US-08-482-073-9	Sequence 9, Appl1
34	100.8	7.5	2861	4	US-08-482-073-10	Sequence 10, Appl1
35	100.8	7.5	3647	1	US-07-914-281-7	Sequence 7, Appl1
36	100.8	7.5	3647	1	US-08-393-246-7	Sequence 7, Appl1
37	100.8	7.5	3647	1	US-08-525-058A-7	Sequence 7, Appl1
38	100.8	7.5	3647	2	US-08-696-731-7	Sequence 7, Appl1
39	100.8	7.5	3647	4	US-09-042-531-7	Sequence 7, Appl1
40	100.8	7.5	3647	5	PCT-US91-00899-4	Sequence 4, Appl1
41	97.6	7.3	2134	2	US-08-483-151-3	Sequence 3, Appl1
42	97.6	7.3	2134	5	PCT-US96-06427-3	Sequence 3, Appl1
43	47	3.5	7218	1	US-08-232-463-14	Sequence 14, Appl1
44	43.2	3.2	1776	1	US-08-722-001-29	Sequence 29, Appl1
45	43.2	3.2	2140	1	US-08-334-698-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-483-151-1
; Sequence 1, Application US/08483151
; Patent No. 5858752

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Holgersson, Jan

TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,151

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Lech, Karen F.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/278001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1814 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-483-151-1

Query Match

Best local Similarity 52.3%; Pred. No. 4, 8e-27;

Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

DB 548 GCTGCGTCTGAGTGTCTACCGAGCGCTGCTGATGCTGTGCTTCCACC 607

QY	595	ACGAGACATCACTGGAGTCTGACAAATTTACCTCAGCAAGCTAAGCCACCTTCCAGA	654
Db	608	ACCGTAGCGTGCAAACCCGGGCAATCTCTCCCTACCCCTGGACACAGAGCACACGAGAC	667
QY	655	AATGGATTTTGGATGAATTTGGAAATCACAACATCACACTCCCAAAAGAGTGGCATTTGAGC	714
Db	668	CTTGGGCTCTGGGCTCCATGGAAATCCGCCAATAATACCATGTCATGCATGCCGG	727
QY	715	ACTGTGTTAATCTGACTCTGACTTACCGCGCGATTACATATATCCAGGCTTATGGCT	774
Db	728	GCATCTTCAACATGGGTGCTGAGCTATCCGCGATTCAGATATCTTTGTACCTTAAGSTC	787
QY	775	TCTTGACGGTAAACACAAATCCCTTGSTGTTTGAAGTGCACGAAGAAAGAAATGGTGT	834
Db	788	GCTTTGAGGCTCTCTCTGCGGC---CACATCCCACATCCGGCCAAAGCAGATGGCTG	844
QY	835	GCTGGGTTTGAGTAACGTGGAACCCCTGAGCATGCCAGATCAATGATTTAAATGAGCTAA	894
Db	845	CTGGGGGATCAGCAATTCACAGAGCGGACAGAGGTGCAAAGGTGATCCGGCAGCGGG	904
QY	895	GCAAAACATTTGAAATTCATACCTACGGGCAAGCAATTTGGAAATATGTCATATGATAAA	954
Db	905	CCCCATCTGCAAGGTGATGTGTTGGTGGCGGCAAGCGAGCGCCCTATGCGCTAATT	964
QY	955	ATTGATTTCTACCATATCTGCTTGTAATTTTATCTTCTTGAATAATTCATCCACA	1014
Db	965	GTCGCGGCCCATTTGGCCCCGGTACCGGTTTCACCTGGCCCTTGAGAACTACAGCATC	1024
QY	1015	AGGATTTACATCAGGAAAGAGCT---ACAATGCTTTTCGGGCGGCTGTACTGTTG	1071
Db	1025	GGGACTCATCATCTGAGAAAGTTCTGGCGCAATGCCCTGGCGGCTGTGCTGATCCGTGG	1084
QY	1072	TTCCTGGACCATCTAGGAAAACCTATGAGAAATATATTTCCAGACATTCATTCATCATG	1131
Db	1085	CGCGGGACCTCTCGGGCCACCTACGAGGCTTTTGTGTGCCACCAATATCCTTTGTACAG	1144
QY	1132	TGGAAGATTATTAATCTCCCGAGTGAAGCTAC	1162
Db	1145	TGAGCAGACTTCAGCTCTGCCCGGAGCTAGC	1175

RESULT 2
 PCT-US96-06427-1
 Sequence 1, Application PC/TUS9606427
 GENERAL INFORMATION:
 APPLICANT: The General Hospital Corporation
 TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/06427
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/483,151
 FILING DATE: 07-JUN-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Lech, Karen F.
 REGISTRATION NUMBER: 35,238
 REFERENCE/DOCKET NUMBER: 00786/278W01
 TELECOMMUNICATION INFORMATION:

```

1      TELEPHONE: 617/542-5070
2      TELEFAX: 617/542-8906
3      TELER: 200154
4      INFORMATION FOR SEQ ID NO: 1:
5      SEQUENCE CHARACTERISTICS:
6          LENGTH: 1614 base pairs
7          TYPE: nucleic acid
8          STRANDEDNESS: single
9          TOPOLOGY: linear
10     MOLECULE TYPE: CDNA
11     CCCT-0596-06427-1:

```

Query Match	9.48;	Score 127;	DB 5;	Length 1814;
Best Local Similarly	52.38;	Pred. No. 4.8e-27;		
Matches 330; Conservative	0;	Mismatches 295;	Indels 6;	Gaps 2;

OY	535	GATCCATGTCACAAAGGACGCTTGTACTGTACAAAGAAATCCATGAGTGTGATGATC	594
Db	548	GCTCCCGCTGATGTCCTAACCCGGAGCGCTGTAACCAAGTGTGATGCTGTGCTTCAC	607
OY	595	ACCCAGACATCAGTGGGATCTGCACAATTTACCTAGCAAGCTAGGCCACCTTTCAGA	654
Db	608	ACCGTGAAGCTGCAAAACCCGGCAATCTCTCTACCCCTGGACAGAGGCCACAGGACAC	667
OY	655	AATGATTTGGATGATTTGGAATTCACCAACTCACACTCCCCAAAAGAGTGGCATTTGAC	714
Db	668	CTTGGGCTGTGGGCTCCATCGAATCCGCCACAGTAATACCATGGTCTCCATCTCCCGG	727
OY	715	ACTTGTTTAACTGATCCTGATCTACTTACCAGCCGCTGATTACATATCCAAAGTCCCTTAGCT	774
Db	728	GCACTCTCAACTGGTGTCTAGCTATACGGGGTATTCAGATATCTTTGACCTTAGGTC	787
OY	775	TCTTGAGGTAAGACAAATCCCTGTGTGTTTAAGTGGCAAGCAAGAAGAAATTTGGT	834
Db	788	GCTTGGAGCCCTCTCTCTGGGCC---CACATCCCACTACCGGCCAAAACAGAGATGGCTG	844
OY	835	GCTGGGTTTGATGATCTAGCAACCTCTGAGCATGCCAGAGTCAAGTATTTGCAATGAGCTAA	894
Db	845	CTGTGGTGATCAGCAATTTCCAGAGAGCGGAGAGCGTCAAAAGCTGACCGGACGTGG	904
OY	895	GCAAAAGCATTTGAATCCATACCTACGGGCAAGCATTTGGAGAAATATGCAATGATATAA	954
Db	905	CCCCCATCTGCAAGTGTGATGTGTGGTCGCCACGAGGAGCGGCCCTATCGCTAATT	964
OY	955	ATTGATTTCCACCATATCTGCTGTGTAATTTATCTTTGAAATTTGAAATTCACCA	1014
Db	965	GTCGTGCGCACTTTGGCCCGGTACCGGCTTACTGGCCCTTTGGAACCTACAGCATC	1024
OY	1015	AGGATTTACATCAGGAAABACTAT---ACAATGCTTTTCTGCGCTGTCTGACTGTG	1071
Db	1025	GGGACTACATACAGAAATTTCTGGCGCATCCCTGGCGGGCTGGTGGTGTACCCGGG	1084
OY	1072	TTCTGGACCATCTAGGAAAACTATGACAATTATTTCCAGACAGATTCATTCATTCATG	1131
Db	1085	CGGTGAGCCCTCTCGGGCCACCTACGAGGCTTTTGTGCCACCAAGTGGCTTTGTACAG	1144
OY	1132	TGGAAGATTATACTCTCCACAGTACAGCTAGC	1162
Db	1145	TGAGACACTTACAGCTGTGCCCTGAACTGCG	1175

```

1  RESULT 3
2  US-07-914-281-13
3  Sequence 13 Application US/07914281
4  Patent No 5324663
5  GENERAL INFORMATION:
6  APPLICANT: LOWE, JOHN B.
7  TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
8  TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
9  TITLE OF INVENTION: GLYCOLIPIDS OR AS FREE MOLECULES AND FOR THE ISOLATION
10 TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
11 NUMBER OF SEQUENCES: 14
12 CORRESPONDENCE ADDRESS:

```

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C. Jefferson Davis Highway, Fourth Floor
 City: Arlington
 State: Virginia
 Country: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/914,281
 FILING DATE: 19920720
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELETYPE: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1654 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-07-914-281-13

Query Match 8.4%; Score 113.4; DB 1; Length 1654;
 Best Local Similarity 51.2%; Pred. No. 4.1e-23;
 Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

462 ATTCTGTGTGGTGGCCATTGGGAGACCTTTGACCTTACATCCCAAGCATG 521
 318 ATCTCTGTGGAGAGTGGCTTTTAAACAACCATACCTGCTGCCCTGAGAGATG 377
 522 TT---CAACATCCAGAGATCCATCTCAACAAGGACCGTCTACCTGTAACAATCCAT 578
 378 GTGCTGTGGAGAGTGGCTTTTAAACAACCATACCTGCTGCCCTGAGAGATG 437
 579 GCAATCTGTATCCATCCAGAGATCCATCTCAACAAGGACCGTCTACCTGTAACAATCCAT 635
 438 GCGGTATCTGTGACCAACGAGAGATCTGTAACAACCGTCTGCTGCCCTGAGAGATG 497
 636 GCTAGGACCACTCTTCAGAAATGATTTGATTAATTCACCACTCACTCC 695
 498 CCGAGGAGGAGGAGGAGGAGATCTGTTGAGATGAGTCCCAAGCAGCTGTGG 557
 696 CAAAAGATGAGATGAGGAGGAGATCTGTTGAGATGAGTCCCAAGCAGCTGTGG 755
 558 CAACTGAAACGAGGAGGAGATCTGTTGAGATGAGTCCCAAGCAGCTGTGG 617
 756 ATCAAGTGGCTTATGAGCT-----TCTTACGAGTAAAGCAAAATCCCTGCTGT 806
 618 ATCTTACGAGCTTATGAGCTGCTGAGAGCTGTGCTGCCAGCTGCCCAAGCAGCTG 677
 807 GAAGTGGCAAGCAAGAAATGATTTGATTAATTCACCACTCACTCC 866
 678 AACCTGCGGCAAGAGAGGAGGAGTGTGGCTGCTGAGATGAGTCCCAAGCAGCTG 737
 867 GCGAGATGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 926
 738 GCCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 794
 927 GCATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 986
 795 CGCTCCCAAGAGCCCTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 854

987 TATCTTCTTTGAAAATTAATCCACAGAGATTACATCAGGAAAAGCTAT---ACAT 1043
 855 TATCTGCTTTCGAGAACTCTTCCACCGACATCAACCGAGAGAGCTGTGAGAGAC 914
 1044 GCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
 915 GCCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
 1104 TATATTCACAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1163
 975 TTCTGTCACACCGAGCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1034
 1164 AAGTATCTGAGAGAGTGCACAAAACATTAAGTTATCTTACTTAACTGAGAG 1223
 1035 CGGTACTGAGAGAGTGCACAAAACATTAAGTTATCTTACTTAACTGAGAG 1094
 1224 AAG 1226
 1095 GAG 1097

RESULT 4
 US-08-393-246-13
 Sequence 13, Application US/08393246
 Patent No. 5595900
 GENERAL INFORMATION:
 APPLICANT: LOWE, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C. Jefferson Davis Highway, Fourth Floor
 City: Arlington
 State: Virginia
 Country: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994
 APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELETYPE: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1654 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-393-246-13

Query Match 8.4%; Score 113.4; DB 1; Length 1654;
 Best Local Similarity 51.2%; Pred. No. 4.1e-23;

Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

```
OY 462 ATTCTGTGGTGGTGGCATTTGGGACACTTGTGACCTTACATCCCGCAAGCATG 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ATCTGTGTGGTGGTGGCATTTTAAACAAACCATAGCTGTGCGTCAGAGATG 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 522 TT---CAACATCCAGAGATGCCATCTACACAGGACCGTTCCTAGTACAAATCCAT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTGCTGGGACGCGTCAATCAATCACTGCGACGAGAGGTGATCCAGAGCGAGC 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 579 GCAGTCTGATCCATCAGGAGATCACTGGGATCTGA---CAATTTACCTGAGCA 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GCGGTATCTGTCACACCGAGAGTGTATCAACACCGAGTCCAGCTCCAGCTCC 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 636 GCTAGGCAACCTTCCAGAAATGGATTTGGATGATGAAATCAACCACTACACTCC 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CCGAGGCGGAGGGGAGCATGATCTGTTCCAGATGAGTCCCAAGCCACTGCTGG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 696 CAAAGAGTGGCATTTGACACCTGTTAAGCTGACTGACTTACCGCGGTATTCAGAT 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CAGCTGAAGCCATGAGAGGATCTTCAATCTCACCATCTCTACCGGACGACCTCCAG 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 756 ATCCAGTGGCTTATGGCT-----TCTGAGGTAAGACAAATCCCTTCTGTTT 806
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 ATCTTACCGCTTACGCTGCTGCGAGCCGTGTCGCGCCAGCTGCGCCACCGCTC 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 807 GAAGTCCCAAGCAAGAAATGGTGTGCTGGGTTGTGATTAAGTGAACCTGAGCAT 866
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AACCTTCCGCAAGACCGAGCTGGTGGCTGGGAGTCCAACTGAGGGGCAAACTCC 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 867 GCCAGAGTCAATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 927 GCATTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 986
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CCGTCCACAAAGCCCTGCGCCAGGAGACCATGATGATGATGATGATGATGATG 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 987 TATCTTCTCTTGAATTAATCAATCCAAAGATTAATCAATCAATCAATCAATCA 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 TATCTGCGCTTGAAGATCTCTGACGCCGACTACATACCGAAGAGTGTGAGGAA 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1044 GCTTTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 915 GCCCTGAGAGCCCTGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1104 TATATTCAGAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 975 TTCTCTGACACCGACGCTTCACTGACGAGAGTTCAGAGAGCCCAAGAGAGCTG 1034
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1164 AAGTATTCAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 CCGTACTCTGAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1224 AAG 1226
    ||
Db 1095 GAG 1097
```

RESULT 5
US-08-525-058A-13
Sequence 13, Application US/08525058A
Patent No. 5770420

GENERAL INFORMATION:

APPLICANT: IOME, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSES: OBLION, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor

```

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,058A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavelle, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)486-2347
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1634 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-525-058A-13
Query Match 8.4%; Score 113.4; DB 1; Length 1634;
Best Local Similarity 51.2%; Pred. No. 4.1e-23;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
```

```
OY 462 ATTCTGTGGTGGTGGCATTTGGGACACTTGTGACCTTACATCCCGCAAGCATG 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ATCTGTGTGGTGGTGGCATTTTAAACAAACCATAGCTGTGCGTCAGAGATG 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 522 TT---CAACATCCAGAGATGCCATCTACACAGGACCGTTCCTAGTACAAATCCAT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GTGCTGGGACGCGTCAATCAATCACTGCGACGAGAGGTGATCCAGAGCGAGC 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 579 GCATTTGATCCATCAGGAGATCACTGGGATCTGA---CAATTTACTGAGCA 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GCGGTATCTGTCACACCGAGAGTGTATCAACCCAGTCCAGCTCCAGCGCTCC 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 636 GCTAGGCCACCTTCCAGAAATGGATTTGGATGATGATGATGATGATGATGATGATG 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 696 CAAAGAGTGGCATTTGACACCTGTTAAGCTGACTGACTTACCGCGGTATTCAGAT 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CAGCTGAAGCCATGAGAGGATTCATTCATTCATTCATTCATTCATTCATTCATTC 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 756 ATCCAGTGGCTTATGGCT-----TCTGAGGTAAGACAAATCCCTTCTGTTT 806
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 ATCTTACCGCTTGAAGATCTCTGACGCCGACTACATACCGAAGAGTGTGAGGAA 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 807 GAAGTCCCAAGCAAGAAATGGTGTGCTGGGTTGTGATTAAGTGAACCTGAGCAT 866
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AACCTTCCGCAAGACCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 867 GCCAGATCAAGTATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 927 GCATTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 986
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CCGTCCCAAGAGCCCTGCGCCAGGAGACCATGATGATGATGATGATGATGATGATG 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 987 TATCTTCTCTTGAATTAATCAATCCAAAGATTAATCAATCAATCAATCAATCA 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 TATCTGCGCTTGAAGATCTCTGACGCCGACTACATACCGAAGAGTGTGAGGAA 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

OY      1044 GCTTTTCGGCGTCCGTGCTACCTGTTGTCGAGCAATCGAATAACATGACAT 1103
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      915 GCCCTGGAGGCCCTGGGCCCTGCCTGGTGCTGGGCCCCAGCAGACGACTGAGAGG 974
OY      1104 TATATTCCAGCAGATTCATTCATCATGTGGAAGATTATTACTGCCAGTAGCTACA 1163
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      975 TTCCTGACCACCCGACCGCTTCATCCACGTCGAGCAGCTTCAGAGCCCAAGAAGACTGGCC 1034
OY      1164 AAGTATCTCGAAGAAAGTGCACAAAACATAATAGTTATACCTTAGTACTTTAGTCGAG 1223
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1035 CGGTACTCGACGACAGCTGACAGAACGACACGCCCGCTACCTGAGAGCTATTCGCTGGCGG 1094
OY      1224 AAG 1226
           ||
DB      1095 GAG 1097

RESULT 6
US-08-696-731-13
; Sequence 13, Application US/08696731
; Patent No.5955347
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUBSTADT,
; ADDRESS: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,731
; FILING DATE: 14-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/393,246
; FILING DATE:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281 -
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-696-731-13

Query Match          8.4%; Score 113.4; DB 2; Length 1654;
Best Local Similarity 51.2%; Pred.No.4.1e-23;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
```

```

OY 462 ATTTGGGTGGGTGGGTCATTGGGGAGACCTTTGACCTTTACATTCCTGGCAAGCATG 521
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 318 ATTCGTCTGTGTGACGTGGCTTTTAAACAACCCATAGCTGTGGCCCGCTGCTCAAGATG 377
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 522 TT--CAACATCCAAAGGATGCCATCTCAACAGGACCGCTTGACTGTACAACAATCCCAT 578
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 378 GTGCGTGGCAGCGGTGACTGCAACATCACTGTGCCAGCCGCAAGGTGTATTCACAGGCACAC 437
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 579 GCATGTTCTGATTCATCCAGACATCACTGGATCTGA--CAAAATTACCTCAGCAA 635
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 438 GGGGTCATCTGTCCACACCGAGAGGTCAATGTAACAACCCAGTGGCCACACTCCACAGCTCC 497
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 636 GCTTAGCCACCCCTTCAGAAATGATTTGGATTTGGATTCAGATCAACAACTCACACTGCC 695
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 498 CCGAGGCGGCGAGGCGCAGATGATCTGTGTTACACATGAGATGCCCAAGCCACTGCTGG 557
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 696 CAAAGAGTGCATTTAGACACTTTTAACTGACTGTACCTGACTTACCGCCGATTCAGAT 755
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 558 CAGCTGAAGAACCATGAGAGGATCTTCAATCTCACCATGTCTCTACCGGAGGACTCGCAC 617
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 756 ATCCAAATGCTTATAGCT-----TCTTACGGTAAAGACCAATTCCTTGTGTTT 806
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 618 ATCTTCACGCCCCCTGAGGCTGGAGCCCTGTGTCTCCGACACCTGCCCACACCGCTC 677
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 807 GAAGTGCCAAAGAGAGAAATTTGTGTGCTGGTGTGAGTAACTGGAAACCTGAGAT 866
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 678 AACCTCTGGGCGCAACACCGAGAGCTGTGGCTGTGGGAGTGTCCAACTGGGGGCAAAATCTC 737
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 867 GCCAGAGTCMAATATTACAAATGAGCTAAAGCAAAACATTTGAATCCATACCTACGGGCAA 926
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 738 GCCAGAGTGTGCGTACTACAGAGCTGCAAGGCCCATCTCAAGGTGAGGTGAGCG--A 794
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 927 GCATTTGGAGATATGTCAATGATATAAAATTTGATTCCTACCATATCTGTCTGTAAATTT 986
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 795 CGCTCCCAACAAACCCCTGCCGCCAGGGAACCAATGATGAGAGCGCTGTCCGGTACAAGTTC 854
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 987 TATCTTTCCTTTGAAATTTCAATCAATCCACAAGATTAATCACTACAGGAATCTAT--ACAAT 1043
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 855 TATCTGGCCTTGAACAACTCTTGTGCACCCCGACTCATCTACCGAGAACTGTGGAGGAC 914
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 1044 GCTTTTGTGGCTGAGCTGTGTACCTGTGTCTTGTGTGGACCATAGGAAAACTATAGAGAT 1103
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 915 GCCGTGAGAGCGCTGGGCCGTGGCCGTGGTGGTGGGCCCCAGAGAAAGCACTAGAGAGG 974
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 1104 TATATTCACAGCATTTCAATCTATGTGTGGAAGATTTAACCTGCCAGTAGAGTACA 1163
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 975 TTTCGCGCACCGCAGCGCTTCACTCACGCTGGACGACTTCCAGAGCCCCAAAGACTGTGCC 1034
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 1164 AAGTATCTTGAAGAGAGTGCACAAAACAATAGTATATCTAGTATCTTAACTGAGGAG 1223
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 1035 CGGTACTCTGACGAGACTGTGACAAAGACACAGCCCGCTACTGTAGCTACTTTCGCTGGCGG 1094
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 1224 AAG 1226
   1 1
DB 1095 GAG 1097

RESULT 7
US-09-042-531-13
; Sequence 13, Application US/09042531
; Patent No. 6268193
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
;

```

Tue Oct 8 10:18:35 2002

us-09-744-748-5_copy_29_1374.rn

Page 6

CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/042,531
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/393,246
 FILING DATE:
 APPLICATION NUMBER: US 08/220,433
 FILING DATE: 30-MAR-1994
 APPLICATION NUMBER: US 07/914,281
 FILING DATE: 20-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)466-2347
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1654 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-09-042-531-13

Query Match	8.48;	Score 113.4;	DB 4;	Length 1654;
Best Local Similarity	51.28;	Pred. No. 4.1e-23;		
Matches 401; Conservative	0;	Mismatches 361;	Indels 21;	Gaps 52;

OY	462	ATTTGTTGGTGTGGCCATTTTGGGAGACCTTTGACTTATCCTGCGCAAGCATG	521
OY	463	TTTGGTGTGGTGTGGCCATTTTGGGAGACCTTTGACTTATCCTGCGCAAGCATG	522
Db	318	ATTCCTGCTGTGGACGGTGGCTTTTAAACAACCCATAAGCTGTGCCCGCTGTGAGAAATG	377
OY	522	TT----CAACATCCAAAGATGCCATCTCACACAGCAGCCCTTACCTGTACCAACAATCCCAT	578
Db	378	GTGGCTTGGCAGGGCTCACTGCAAACTACATCGCTGCCGCCAGAGGTGTATCCACAGGGCAGAC	437
OY	579	GCATTTGTGATTCATCACCGACAGCATCATGATGGATTCGA---CAAAATTACCTCAGCAAC	635
Db	438	GGGTGATCATGTGTGACCAACCGAGAGGTGATGTACAAACCCATGTGCCAGCTGCCACGTCC	497
OY	636	GCTAGGCGACCTTCCAGAAATGSGATTTGATGAAATTGGATATCAACACTCACACTCCC	695
Db	498	CCGAGGGGAGGGGACAGCATGAGATCTGTGTACACATGAGTATCCCAAGCACTGTGTGG	557
OY	696	CAAAAGATGGCACTTAGCACTTGTTTTAAACCTGACTGTGACTTACCGCCGTGATTCAGAT	755
Db	558	CAGCTGAACCACTGACAGCGATTACTTAATCTCACCAATGCTCTACCGCAGAGACTCGAC	617
OY	756	ATCCAAATGCTTATGCT-----TCGTACAGGTGAACACAAATTCCTTGATGT	806
Db	618	ATCTTACAGCGCTTACGCGCTGGCTGGAGCGGTGTTCGGCCAGCCTGCCCAACCAACCGCTC	677
OY	807	GAAATGCCAACAAGAAATTTGGTGTGTCTGGTGTGTGAGTAATCTGGAACCTGAGCAT	866
Db	678	AACCTCTGGCGCAAGACCGAGTGGTGGGTGGGAGATGTCTCAACTTGGGGGCCAAATCTCC	737
OY	867	GCCAGACTCAAGATTACAAATGAGTAAAGCAAAAGCACTTGAATTCATACCTACAGGGCAAC	926
Db	738	GCCAGGGTGTGGCTACTACAGAGCCCTGACAGGCCATCTCAAGGTGTGAGAGGTGAAGG---A	794

OY	927	GCAATTGGAAATATGTCATATGATATAAATTTGATTCCTACCAATATCTGCTTGTAATTT	966
Db	795	GCGTCCACACAGCCCCCTGCCCCAGGAAACCATGATGGAGACGCTGTGCCGGTACAAAGTTTC	854
OY	987	TATCTTTTCTTTGAAATATTCATATCCACAGAGATTACATACGAGAAAAGCTAT---ACAAT	104
Db	855	TATCTGGCTCTGGAGAACTCTTGACCCGACCTACATATACGAGAAAGCTCTGGAGAAC	914
OY	1044	GCTTTTCTGGCTGCTCTGTACTGTGTTGTTCTGGACCATCTAGGAGAAACTATAGAAAT	110
Db	915	GCCCTGGAGCCCTGGGCGCCCTGCTGCTGTGGGCCCAAGCAAGAAAGCAACTCTAGGAAGG	974
OY	1104	TATATTCACAGAAATTCATTCATTCATGTGGAAATTTATTACTCTCCAGTAGAGCTAGCA	116
Db	975	TTCCTCGCACCCGACGCGCTTATCATCCATGTGAAGACTTCCAAAGGCCAAAGAACTGTGCC	103
OY	1164	AAGTATCTGAAGGAATGGACAAAAACAATTAAGTTTATACCTTATAGTTTAACTTGAGG	1222
Db	1035	CGGTACTCTAGAGACTGGACAAGGACCGCCGCTTACCTGATAGCTTCTTGCTTGCGGG	109
OY	1224	AAG 1226	
Db	1095	GAG 1097	

RESULT 8
US-07-914-281-12
; Sequence 12, Application US/07914281

GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C.
ADDRESS: 1755 Jefferson Davis Highway, Fourth Floor
City: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ivalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-914-281-12

Query Match	8.3%;	Score 112.2;	DB 1;	Length 1086;
Best Local Similarity	51.6%;	Pred. No. 7e-23;		
Matches 364; Conservative	0;	Mismatches 323;	Indels 18;	Gaps 4;

OY	537	TGCCCTCCACAAAGGACCCTTCACTGACACAAATGCCATGCACTCTGATCCATCAC	596
Db	271	TGCCACATCACTGCGCGACCGGAGAGGTGATCCACAGGACGACAGCGTATGTCGACACC	330
OY	597	CGAGACATCACTGTGGGA---TCTGACAAATTTACTCTGACGAAGCTAGGCGACCCCTCCAG	653
Db	331	TGGGATATCATGCTCCAAACCCTTAAGTCACGCTCCACCTTCCCGAGGCGCGAGGGCGAG	390
OY	654	AAATGGATTTGGATGAAATTTGGAAATCACCAACTGACACATCCCCAAAAGAGTGCAATTGAG	713
Db	391	CGCTGGATCTGGTTCAACTTGGAGCAGCCCTTAAGTCCGACACCTGGAAGCCCTGGAC	450
OY	714	CACTTGTTAACTGACTCTGACTTACCGCCGTGATTCAGATATCCAAAGTCCCTTAGGC	773
Db	451	AGATACCTCAATCTCACCATGTCTCTACCGCAGCGACTCGACATCTTCCACCCCTACGCG	510
OY	774	TTCTTGACGGTATGACAAATATCCCTTGGT-----GTTTGAAGTCCCAAGCAAAAG	824
Db	511	TGCGTGGAGCCGTGGTCCGCGCACGCTGCCACCCACCCGCTCAACTCTCTGGCCAAAGACC	570
OY	825	AAATTTGGTGTGCTGGGTTGTGAGTAACTGGAACCCCTGAGATGCGAGATCAATATATAC	884
Db	571	GAGCTGGGTGGCTGGGGGTGTCCAACTGGAAAGCCGGAGCTACGACGAGGTGCGCTACTAC	630
OY	885	AATGAGCTAAGCAAAAGCAATTGAATCCATACCTACGGGGCAGCAATTTGGAAATATGTC	944
Db	631	CAGAGCCCTGAGGCTCATCTTCAGAGTGAGACGTGATGAGACGCTCCACAAACCCCTGGCC	690
OY	945	AATATATAAAATTTGATTTCCATCATCTGCTTGTAAATTTATCTTTCTTCTTGGAAAT	1004
Db	691	AAGGGGACCA---TGATGGAGACGCTGTCCCGGTACAAAGTCTACCTGGCCCTTGAGAAC	747
OY	1005	TCAATTCACAGAAATATACATCAACGAAAAGCTAT---ACAATGCTTTCTGCTGGGCTCT	1061
Db	748	TCTCTTGACACCCCGACTACATCAACGAGAAAGCTGTGGAGAAACGCCCTTGAAGCCTGGGCC	807
OY	1062	GTAACCTGTTGTTCTGGGACATCTAGGGAACAACTATGAAATTAATTCACGAGATTTCA	1121
Db	808	GTGCGCCGTTGTTGGGGCCCGACGAGAAAGCAATACGAGAGAGTTCTCCACCCGACGCC	867
OY	1122	TTCAATTCATGTGGAAGATTAATTAATCTCCAGTAGACTAGCAAAAGTATCTGAAGAAATC	1181
Db	868	TTTCATCCACGTGGAGCACTTCCAGAGCCCCCAAGGAACTGGCGCCGTACTCTCAGAGAGCTG	927
OY	1182	GACAAAAACAATATAGTTATACCTTAGTTACTTAACTGAGAGAAAG	1226
Db	928	GACAAAGGACACGCCCGCTACTAGCTACTTTCGCTGGCGGAG	972

RESULT 9
 US-08-393-246-12
 Sequence 12 Application US/08393246
 Patent No. 5593500
 GENERAL INFORMATION:
 APPLICANT: LOWE, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-393-246-12

```

Query Match 8.3%; Score 112.2; DB 1; Length 1086;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY	537	TGGCATTCACCAAGGCGGTGCATAGTAAACAATAATCCATGCACTTGATCATTAC	596
Db	271	TGCCACATCACTGCGGACCAGGAGTGTACCCACAGGCAAGACAGGCTATGTGGACAC	330
OY	597	CGAGACATCAAGTTGGGA--TCTGACAAATTTACTCTACAGAAGTAGGCGACCTTCCAG	653
Db	331	TGGGATATCAATGTCCAACCTTAAGTACAGCGCTCCACACTTCCCOCGAGGCCGAGGGGAG	390
OY	654	AAATGATTTTGGATGAATTTTGGAAATCACCACTCACACATCCCCAAAAGATGGCATTTGAG	713
Db	391	CCTGGGATCTGGTTCAACTTGGAGCAACCCCTTAOCTGGCAGACCTGGAMACCTCTGGAC	450
OY	714	CACITGTTTAACTGACTCTGACTTACCGCGCGGATTCAGATATCCAAAGTCCATTAGC	773
Db	451	AGATACTTCATCTCACACATGTCTCTACGCCGAGGAGATCCGACATCTTACGCCCTTACGCC	510
OY	774	TTCTTACGCTAAGCACAAATCCCTTGCT-----GTTTAAAGTGGCAAAGCAAGAG	824
Db	511	TGGCTGGAGCGGTGTCTCGGCGACGCTGCCACCCACCCCTCAACCTCTCGGCACAAAGCC	570
OY	825	AAATGTGTGTGCGGTTGTGTGAGTAACGTGAACCTCGAGCATGCGCAGATTCAAAGTATAC	884
Db	571	GAGTGGTGGGCTGTGGGGGTGTCCAACTGGAAGCCGACTACAGCCAGGTTGGCTTACTAC	630
OY	885	AATGAGTAAAGCAAAAGCATTTGAATTCATACCTACGGGCAAGCATTTGGAAATATGTC	944
Db	631	CAGAGCCTCGAGGCTCATCTCAAGGTGGAGCTGTACGAGCGCTCCACAAAGCCCTGCCC	690
OY	945	AATATATAAATTTGAATTCCTACCATATGCGTGTAAATTTTATCTTCTTTAAAAAT	1004
Db	691	AAGGGGACCA---TGATGGAGAGCGTGTCCCGGTACAAAGTTCTTACTGGCTGTGAGAAC	747
OY	1005	TCAATCCACAAGATTACATCACGGAAAAGCTAT--ACAATGCTTTTCTGTGCTGGCTCT	1061
Db	748	TCTCTGACCCCGCATCATCACCGAAGAGCTGTGGAGGAAGCCCTGGAGGCTGTGGGCG	807
OY	1062	GTAACCTGTCTCTGGGAGCACTATAGGAAAATATGAATAATTAATTTCTCAGCAATTTCA	1122
Db	808	GTGCCCCGTGGTGGGCCCCACAGCAAGCAACATACGAAGAGTCTCTGTGCACCCAGCCGCC	867
OY	1122	TTCTATTATGTGAAGTTATTAACCTGCCAGTAGCACTACGAAAGTATCTGAAAGAGATC	1181

Db	868	TTCTATCCACAGCTGGACGACTTTCGACAGGGCCCAAGAGACCTGGCCCGGTACTCTGACAGAGCTG	927
Oy	1182	GACAAAAACATTAAGTTTATACCTTACTTACTTTTACTGGAGAG	1226
Db	928	GACATGGACACAGCCCGCTACTCTAGCTACTTTGCTGGCGAGAG	972

RESULT 10
US-08-525-058A-12
; Sequence 12, Application US/08525058A
; Dataset No. 5770420

APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

ADDRESS: OBLON, SPIVA, MCCLELLAND, MAIER & NEUSS
STREET: 1755 Jefferson Davis Highway, Fourth Floor
City: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATE: 05/08/525, 058A
APPLICATION NUMBER: 05/08/525, 058A

```

1 CLASSIFICATION: 435
2
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Lavalleye, Jean-Paul M. P.
5 REGISTRATION NUMBER: 31,451
6 REFERENCE/DOCKET NUMBER: 2353-060-55
7
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: (703)521-4500
10 TELEFAX: (703)486-2347
11
12 TELE: 24855 OPAT UR
13
14 INFORMATION FOR SEQ ID NO: 12:
15
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 1086 base pairs
18 TYPE: nucleic acid
19 STRANDEDNESS: double
20 TOPOLOGY: linear
21
22 MOLECULE TYPE: DNA (genomic)
23
24 OS-08-525-058A-12

```

Query Match	8.3%	Score 112.2	DB 1	Length 1086
Best Local Similarity	51.6%	Pred. No. 7e-23		
Matches 364	Conservative 0	Mismatches 323	Indels 18	Gaps 4

[illegible]

Db	511	TGCGTGGAGGCCCTGTCGTCCGGCCAGCCCTGCCACCACCCGCTCAACCTCTCCGGCCAAAGAC	570
OY	825	AAATTTGGTGTGCTGCTGGGTTTGTCGAGTAACTGGAAACCTTGAGCATGCCAGAGTCAAGTATTAAC	884
Db	571	GAGCTGGTGTGCTGGGGGGTGTCCAACTGGAAACCCGACTCAACAGGGGTGCGTACTAC	630
OY	885	AATGAGCTAAGCAAAAGACATTGGAATTCATACCTACGGGCCAAGCATTTGGAGATATATGTC	944
Db	631	CAGAGCCTCGAGGCTCATCTCAAGGTGGAGCTGTACGGAGCGCTCCCAAGGCCCTGTGCC	690
OY	945	AATGATAAATTTGATTTCTTACCATATCTGCTTGTAATTTATCTTCTTTGAAAT	1004
Db	691	AAGGGAGCA---TGATGGAGAGCGTGTGCCGGTACAGATTTCACTGGGCTCTCGAGAAC	747
OY	1005	TCGATCCACAAAGATTTCATCAGACAGGAAAGACT---ACAAATGCTTTTGGCTGGCTCT	1065
Db	748	TCTTCGACCCCGACTACATACCGAAMACTGTGGAGAACGCCCTGGAGGCTGAGGCC	807
OY	1062	GTACCTGTGTTCTTGGGAGCAATCTAGGGAAACTATGAGATTTATATTTCCACAGATTCA	1121
Db	808	GTGCGCCGTGGTGGCGGCCCCAGCAGAAAGCACTACGAGAGGTTCCTGCGACCCGACGCTC	867
OY	1122	TTCAATTATGTGGAAAGTTTAACTCTCCCAAGTACGCTAGCAAAAGTATCGAAGAGTGC	1181
Db	868	TTTACTCCACGTGGAGACGACTTCCAGAGGCCCAAGAGACTGGCCCGGTACTCTGACGAGAGCTG	927
OY	1182	GACAAAACCAATAGTATTAACCTTAGTACTTTTAATCTGAGGAAG	1226
Db	928	GACAGAGACCAAGCCGCGTACGTAGGTATTTGCTGGGGGGAG	972

RESULT 11
US-08-696-731-12
; Sequence 12, Application US/08696731

1 APPLICANT: LOWE, JOHN B.
2
3 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
4
5 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
6
7 TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
8
9 TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
10
11 NUMBER OF SEQUENCES: 14
12
13 CORRESPONDENCE ADDRESS: 14

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1 APPLICATION NUMBER: US/08/6936/731
2 FILING DATE: 14-AUG-1996
3 CLASSIFICATION: 435
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: 08/393,246
6 FILING DATE:
7 APPLICATION NUMBER: US 08/220,433
8 FILING DATE: 30-MAR-1994
9 APPLICATION NUMBER: US 07/914,281
10 FILING DATE: 20-JUL-1992
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Lavalleye, Jean-Paul M. P.
13 REGISTRATION NUMBER: 31,451
14 REFERENCE/DOCKET NUMBER: 2363-060-555
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (703)521-4500
17 TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-696-731-12

Query Match
Best Local Similarity 51.6%; Pred. No. 7e-23;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

8.3%; Score 112.2; DB 2; Length 1086;
537 TGCATCTCAGACGACCGCTTACTGATACAAATCCAGTGTGATCATCAC 596
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
271 TGCACATCATGCGCCAGAGGTGATCCACAGCAGACGATCATGTGTCACAC 330
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
597 CGAGACATCATGTTGGG---TCTGACAAATTTACCTCAGCAAGTACGCCCTTCAG 653
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
331 TGGATATCATGTCACACCCCTAAGTCCAGGCTCCACCTTCCCGAGGCGGAG 390
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
654 AATGATTTGGATGATTTGGATTCACCACTCAGACTCCCAAAAGAGGCGATTGAG 713
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
391 CGCTGATCTGTGTCACACTTGAGACCCCTTAAGTCCAGCCTGGAAGCCCTGAG 450
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
714 CACTTGTAACTGACTGACTTACCGCGGTGATTGAGATTCAGATGCTTATGCG 773
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
451 AGATCTTCAATCTCACCATCTCTACCGAGCGATCGCATCTTCCAGCCCTAGGCG 510
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
774 TTTCTGACGATGACACAAATCCTTGT-----GTTGAAGTCCCAAGAAAGAG 824
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
511 TGGCTGAGCGGTGTGTCGGGCGACCTGCGCCACCCGCTCAACCTCTCGGCCAG 570
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
825 AATTTGCTGCTGGTGTGATGAGTGAACCTGAGCATGCGAGATCAATATTAC 884
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
571 GAGCTGTGCTGTGGCGGTGTCCAACTGAGAGCGGATCTGAGCGAGTGGCTACTAC 630
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
885 AATGAGCTAAGCAAAAGCATTTGAATTCATCTACGCGGCAAGCATTTGAGAAATATGTC 944
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
631 CAGAGCCTGACAGGCTCATCTCAAGGTGAGGTGAGGAGGCTCCCAAGCCCTGCGC 690
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
945 AATATATATAATTTGATTCCTACCATATCTGCTTGAATTTATCTTCTTGAATAT 1004
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
691 AAGGGAGCA---TGATGAGACCTGTCCGCTGACAAAGTCTACCTGCGCTTGAGAAC 747
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
1005 TCAATCCACAAGATTCATACAGGAAAGCTAT---ACAATGCTTTCTGCTGGGCT 1061
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
748 TCTTTCACACCCCGACTACATCAGAGAGCTGTGAGAGAGCGCTGAGAGGCTTGGCC 807
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
808 GTGCGCTGTGGTGTGGGCCCCAGCAAGAGCACTAGAGAGTCTGCGCACCGGAGGCC 867
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
1122 TTTATTTATGTGGAAGTTTAACTCTCCAGTGAAGTACGAAAGTCTGGAAGAAATC 1181
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
868 TTTATTCACGTGAGCACTTCCAGAGGCCCAAGAGACTGCGCCGCTACCTGACGAGAGCTG 927
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
1182 GACAAAAACAATAATGTTATACCTAGTACTTTAACTGGAGGAG 1226
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
928 GACAAAGACACGCCCGCTACCTGACTTTCGCTGGCGGAG 972
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

RESULT 12
US-09-042-531-12
Sequence 12, Application US/09042531
Patent No. 6268193
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavelle, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-12

Query Match
Best Local Similarity 51.6%; Pred. No. 7e-23;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

8.3%; Score 112.2; DB 4; Length 1086;
537 TGCATCTCAGACGACCGCTTACTGATACAAATCCAGTGTGATCATCAC 596
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
271 TGCACATCATGCGCCAGAGGTGATCCACAGCAGACGATCATGTGTCACAC 330
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
597 CGAGACATCATGTTGGG---TCTGACAAATTTACCTCAGCAAGTACGCCCTTCAG 653
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
331 TGGATATCATGTCACACCCCTAAGTCCAGGCTCCACCTTCCCGAGGCGGAG 390
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
654 AATGATTTGGATGATTTGGATTCACCACTCAGACTCCCAAAAGAGTGGCGATTGAG 713
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
391 CGCTGATCTGTGTCACACTTGAGACCCCTTAAGTCCAGCCTGGAAGCCCTGAG 450
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
714 CACTTGTAACTGACTGACTTACCGCGGTGATTGAGATTCAGATGCTTATGCG 773
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
451 AGATCTTCAATCTCACCATCTCTACCGAGCGATCTGCAATCTTCCAGCCCTAGGCG 510
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
774 TTTCTGACGATGACACAAATCCTTGT-----GTTGAAGTCCCAAGAAAGAG 824
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
511 TGGCTGAGCGGTGTGTCGGGCGACCTGCGCCACCCGCTCAACCTCTCGGCCAG 570
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
825 AATTTGCTGCTGGTGTGATGAGTGAACCTGAGCATGCGAGATCAATATTAC 884
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
571 GAGCTGTGCTGTGGCGGTGTCCAACTGAGAGCGGATCTGAGCGAGTGGCTACTAC 630
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
885 AATGAGCTAAGCAAAAGCATTTGAATTCATCTACGCGGCAAGCATTTGAGAAATATGTC 944
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
631 CAGAGCCTGACAGGCTCATCTCAAGGTGAGGTGAGGAGGCTCCCAAGCCCTGCGC 690
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

OY 945 AATGATATAATTTGATCTTACATATCTGCTTGAATTTTATCTTCTTGAAT 1004
DB 691 AAGGGAGCA---TGATGAGACGCTGTCCGCTACAAAGTCTCTGCTGAGAAC 747
OY 1005 TCAATTCACAGGATTAATCAACGAAAGCTAT--ACAATGCTTTTGGCTGCTCT 1061
DB 748 TCCCTGACACCCGACTATCATCCGAGAGAGCTGTGAGAGAGCCCTGAGGCC 807
OY 1062 GTACTGTGTCTGAGACATCTAGGAAACATATGAAATTAATTCACAGCATTTCA 1121
DB 808 GTGCCCCGTGTGCTGAGCCCGCAGAGAACATCACTAGAGAGTCTCTGCCACCCAGCC 867
OY 1122 TTCATTCATGTGAGATTAATTAATCTCCAGTGAAGTGAAGTATCTGAAGAGTTC 1181
DB 868 TTCATTCACGTGAGAGCTTCCAGAGCCCGCAGAGAGCTGTGCCCCGTGCTGAGAGAGCTG 927
OY 1182 GACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1226
DB 928 GACAAAGACACGCGCCGCTACTGAGTACTTTCGCTGCGGAG 972

RESULT 13

US-09-063-237-2
Sequence 2, Application US/09063237
Patent No. 6124267
GENERAL INFORMATION:
APPLICANT: McEever, Rodger P.
TITLE OF INVENTION: O-glycan Inhibitors of Selectin Mediated
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30306-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/063,237
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/649,802
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF110CIP7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-063-237-2

Query Match 8.3%; Score 112.2; DB 3; Length 2042;
Best Local Similarity 51.6%; Fred. No. 1e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;
OY 537 TGCCATTCACAGGACGCTTCACTGTACAAACAAATCCCATGAGTCTGATCCATCAC 596

DB 343 TGCCATTCACAGGACGCTTCACTGTACAAACAAATCCCATGAGTCTGATCCATCAC 402
OY 597 CGACATCATGTTGGGA---TGTGACAAATTAATCTAGCAAGCTAGGCCACCTTCAG 653
DB 403 TGGATTAATCATGTCACACCCCTTAAGTACAGGCTCCACCTTCCCGGCGGAGGAGGAG 462
OY 654 AATGATTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 713
DB 463 CGCTGATCTGTGTTCACTTGAAGCCACCCCTTAATGTCAGAGACCTGAGAGCCCTGAG 522
OY 714 CACTGTTTAACCTGACTGTGACTTACCGCGGATTCAGATTAATTAATTAATTAATTAAT 773
DB 523 AGATTAATTCATGTCATCATGTCCTTACCGGAGGACTCCGACATCTTCACGCTTACGCG 582
OY 774 TTTGTGACGGTAAAGCAAAATCCCTTCGT-----GTTTGAAGTGGCAAGCAAGAG 824
DB 583 TGGCTGAGCGCTGTGTCGGGCAAGCCCTGCCCACCAAGCTCAAGCTTCGCGCAAGAG 642
OY 825 AATTTGCTGTGCGGTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 884
DB 643 GAGCTGTGCGCTGTGCGGTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 702
OY 885 AATGAGCTTAAGCAAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 944
DB 703 CAGAGCTTGAAGCTCATTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 762
OY 945 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1004
DB 763 AAGGGAGCA---TGATGAGAGAGCTGTCCGCTGACAGTCTTACCTGCTGCGAGAAC 819
OY 1005 TCAATCCACAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1061
DB 820 TCCCTGACCCCGACTCATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
OY 1062 GTACTGTGTCTGAGACCATTAAGGAAACATTAATTAATTAATTAATTAATTAATTAATTA 1121
DB 880 GTGCCGTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 939
OY 1122 TTCATTCATGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1181
DB 940 TTCATTCACGTGAGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
OY 1182 GACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1226
DB 1000 GACAAAGACACGCGCCGCTACTGAGTACTTTCGCTGCGGAG 1044

RESULT 14

US-07-914-281-1
Sequence 1, Application US/07914281
Patent No. 5324663
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATIER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavallee, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ANTI-SENSE: NO
US-07-914-281-1

Query Match 8.3%; Score 112.2; DB 1; Length 2043;
Best Local Similarity 51.6%; Pred. No. 1e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 537 TGCATCTCACACGACCGCTTCACTGACAAATCCATGCAAGTTGATCCATCAC 596
DB 343 TGCACATCACTGCGACCGCAAGGTGTACCCACAGCAGACGGTCATGTCACAC 402
QY 597 CGAGCATCAGTTGGGA---TCTGACAAATTTACTCAGCAAGTAGCCACCTTCAG 653
DB 403 TGGATATCACTGTCACACCTTAAGTCAAGCTCCACCTCCGAGGCGGAG 462
QY 654 AATGATTTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 713
DB 463 CGGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
QY 714 CACTTGTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 773
DB 523 AGATACCTCAATCTCACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 582
QY 774 TTTGAGCGTAAAGCAAAATCCCTGCT-----GTTTGAAGTCCCAAGCAAGAG 824
DB 583 TGGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
QY 825 AATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
DB 643 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
QY 885 AATGAGCTTAAGCAAAAGCAATTTGAATTCATCTAGCGGCAAGCATTGGAATATGTC 944
DB 703 CAGAGCTGCAAGCTCAATCTCAAGTGTGAGCTGACGACCTCCCAACCCCTGCCC 762
QY 945 AATGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1004
DB 763 AAGGGGACCA---TGATGAGACGCTGCTCCCGTAAAGTCTTCTGCTGCTGCTGCTG 819
QY 1005 TCAATCCACAAGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1061
DB 820 TCCCTGACACCGGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 879
QY 1062 GTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121
DB 880 GTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
QY 1122 TTTGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1181
DB 940 TTTGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 999
QY 1182 GACAAAAACAATAAGTTATACCTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 1226
DB 1000 GACAAAGACACGCGCTACCTAGCTACTTTCGCTGCGGAG 1044

RESULT 15
US-08-393-246-1
Sequence 1, Application US/08393246
Patent No. 5595900
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavallee, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ANTI-SENSE: NO
US-08-393-246-1

Query Match 8.3%; Score 112.2; DB 1; Length 2043;
Best Local Similarity 51.6%; Pred. No. 1e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 537 TGCATCTCACACGACCGCTTCACTGACAAATCCATGCAAGTTGATCCATCAC 596
DB 343 TGCACATCACTGCGACCGCAAGGTGTACCCACAGCAGACGGTCATGTCACAC 402
QY 597 CGAGCATCAGTTGGGA---TCTGACAAATTTACTCAGCAAGTAGCCACCTTCAG 653
DB 403 TGGATATCACTGTCACACCTTAAGTCAAGCTCCACCTCCGAGGCGGAG 462
QY 654 AATGATTTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 713
DB 463 CGGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522
QY 714 CACTTGTAACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 773
DB 523 AGATACCTCAATCTCACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 582
QY 774 TTTGAGCGTAAAGCAAAATCCCTGCT-----GTTTGAAGTCCCAAGCAAGAG 824

Tue Oct 8 10:18:35 2002

us-09-744-748-5_copy_29_1374.rni

Page 12

```
Db 583 TGGCTGGAGCCGCTGTCGCCGCCAGCCTGCCACCACCCGCTCAACCTCTCGGCCAAGACC 642
OY 825 AATTGGTGTGCTGGTGTGTGAGTAACGTGAGACCTGAGCANTSCCAGAGTCAAGATTTAC 884
Db 643 GAGCTGTGGCGCTGGCGGTGTCCTCAACTGGAACCCGAGCTCAGCCAGGGTGCCTACTAC 702
OY 885 AATGAGCTAAGCAAAAGCATTTGAATCCATACCTACGGGCAAGCATTGTGAGAATATGTC 944
Db 703 CAGAGCCTGAGGCTCATCTCAAGGTGAGGTGACGTGACGCTCCACAAGCCCTGCCC 762
OY 945 AATGATAAATAATTGATTCCTACCATATCTGCTTGAATAATTTATCTTCTTTGAAAT 1004
Db 763 AAGGGGACCA---TGATGGAGAGCGTGTCCCGGTACAAAGTTTACCTGGGCTTCGAGAAC 819
OY 1005 TCAATCCACAAAGGATTACATCAGGAAAGCTAT---ACAATGCTTTTGTGGGTGCTCT 1061
Db 820 TCCTTGACCCCCGACTACATCACGAGAAAGCTGTGAGAGAACGCCCTGGAGGCTGGGCC 879
OY 1062 GTACCTGTGTCTGTGGAGCACTAGGAAATATGAGAAATATATTCAGAGATTC 1121
Db 880 GTGCCCGTGTGTGTGGGCCCCAGAGCACTACGAGAGTTCCTGCCACCGAGGCC 939
OY 1122 TTCATTCATGTGAGAAATATATACCTCTCCAGTGTAGCAAAAGTATCTGAAGAGATC 1181
Db 940 TTCATTCACGTGTGAGCACTTCCAGAGCCCCAAGAGACCTGGCCGGTACTGTGCAGAGCTG 999
OY 1182 GACAAAACATATAGTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1226
Db 1000 GACAAAGGACACGCCCGCTACCTGAGCTACTTTCGCTGGCGGGAG 1044
```

Search completed: October 6, 2002, 07:22:15
Job time : 39.4635 secs

JOURNAL Submitted (29-JAN-1999) Mika Kaneko, Institute of Life Science,
Soka University, Division of Cell Biology, Soka University, 1-236
Tengai-cho, Hachioji, Tokyo 192-8577, Japan
(E-mail:mike@soc.soka.ac.jp, Tel:81-426-91-2495(ex.5132),
Fax:81-426-91-915)

FEATURES

SOURCE

1. 3019
/organism="Homo sapiens"
/db_xref="taxon:9606"
295..1374
/gene="FUR9"
295..1374
/gene="FUR9"
/codon_start=1
/product="alpha-1,3-fucosyltransferase IX"
/protein_id="BA81685.1"
/translation="MTSTSKGILRPFLIVCIIIGCFMACILYIKPNSWIPSPMSA
SSVILKKNPESTKEDYFNERTILVWVPRGCTDLSCOMELGCHLTTPDSYRK
SHAVLIHHRIISMDLNLROARBPPOKIMMLRSPHPRKSGIEHFNILTYRR
DSDIQYRGITVSTNPVPEVPSKELKLVYVSNNEPBARIKYINELSKSEIHTT
GAPRGVYDNKNIIPITISACKYTLSENSIHKDYIEKLINATLAVSVVLPDSRN
YENIIPDSRIHYEDINSPBELAKTLEVDKNKLTLSIFNMKRDFTVNLPRWESHA
CLACDHYKHEKSYGNLEKPMW"

CDS

BASE COUNT 937 a 594 c 541 g 947 t
ORIGIN

Query Match 100.0%; Score 915; DB 9; Length 3019;
Best Local Similarity 100.0%; Pired. No. 8.6e-231;
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAACGTATTTTATGAACACTATCTGCTGCTGGCTGGCCATTGGGCAG 60
DB 460 ACCAAACGTATTTTATGAACACTATCTGCTGCTGGCTGGCCATTGGGCAG 519
QY 61 ACCTTTGACCTTACATCCTGCGCAAGCAATGTCACATCCAAAGATGCATCTCACAAG 120
DB 520 ACCTTTGACCTTACATCCTGCGCAAGCAATGTCACATCCAAAGATGCATCTCACAAG 579
QY 121 GACGCTCAGCTGTACACAAATCCCATGCTGATCCATCCGAGAGCATAGTTGG 180
DB 580 GACGCTCAGCTGTACACAAATCCCATGCTGATCCATCCGAGAGCATAGTTGG 639
QY 181 GATCTGACAAATTTACTGACGACGATAGCCACCTTCCGAAATGATTTGGATGAT 240
DB 640 GATCTGACAAATTTACTGACGACGATAGCCACCTTCCGAAATGATTTGGATGAT 699
QY 241 TTGGAATCACCACCTACACCTCCCAAAAGAGTGCAATGACACTGTTTAACTGACT 300
DB 700 TTGGAATCACCACCTACACCTCCCAAAAGAGTGCAATGACACTGTTTAACTGACT 759
QY 301 CTGACTACCGCCGCTGATTCAGATATCCAGTGCTTATGCTTCTGACGGTAAGCACA 360
DB 760 CTGACTACCGCCGCTGATTCAGATATCCAGTGCTTATGCTTCTGACGGTAAGCACA 819
QY 361 AATCCCTTCTGCTTGAAGTGCCAAAGCAAGAAATTTGGTCTGGTGTGAGTAAC 420
DB 820 AATCCCTTCTGCTTGAAGTGCCAAAGCAAGAAATTTGGTCTGGTGTGAGTAAC 879
QY 421 TTGGAACCTGAGCATGCGACAGTGCAAGTATTTACAAATGAGTAAAGCAAGATGAATC 480
DB 880 TTGGAACCTGAGCATGCGACAGTGCAAGTATTTACAAATGAGTAAAGCAAGATGAATC 939
QY 481 CATACCTACGCGGCAAGCATTTGGAGATATGTCATATAAAATTTGATTCCTACATA 540
DB 940 CATACCTACGCGGCAAGCATTTGGAGATATGTCATATAAAATTTGATTCCTACATA 999
QY 541 TCTGCTTGTAAATTTTATCTTCTTGGAAATTCATCCAAAGATTTACATCAGCGAA 600
DB 1000 TCTGCTTGTAAATTTTATCTTCTTGGAAATTCATCCAAAGATTTACATCAGCGAA 1059
QY 601 AAGCTATCAATGCTTTCTGGCTGGCTGTGACTCTTGTCTGGAGCATCTAGAGAA 660
|||||

DB 1060 AAGCTATCAATGCTTTTCTGGCTGGCTGTGACTCTGTCTGCGAGCAACCTAGCGAA 1119
QY 661 AACTAGCAATTTATATTCACAGCATTCATTCATTCATGAGATTTAACTCTCC 720
DB 1120 AACTAGCAATTTATATTCACAGCATTCATTCATTCATGAGATTTAACTCTCC 1179
QY 721 AGTCAGCTACCAAGATATCTGAAAGAGTGACAAAAACATTAAGTTAACTTACTTAC 780
DB 1180 AGTCAGCTACCAAGATATCTGAAAGAGTGACAAAAACATTAAGTTAACTTACTTAC 1239
QY 781 TTTAACTGAGGAAGATTTCTGTAATCTTCCAGATTTTGGAGATCAGATGATGT 840
DB 1240 TTTAACTGAGGAAGATTTCTGTAATCTTCCAGATTTTGGAGATCAGATGATGT 1299
QY 841 TTGGCTTGCATCATGTAAGAAAGCATCAAGATTAATTAAGTCTGTGTAATTTAGGAAA 900
DB 1300 TTGGCTTGCATCATGTAAGAAAGCATCAAGATTAATTAAGTCTGTGTAATTTAGGAAA 1359
QY 901 TGGCTTTGGAATTTA 915
DB 1360 TGGCTTTGGAATTTA 1374
|||||

RESULT 2
AL512406
LOCUS
DEFINITION Human DNA sequence from clone RP11-504J9 on chromosome 6, complete
sequence.
ACCESSION AL512406
VERSION AL512406.14 GI:13897154
KEYWORDS
SOURCE
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (27-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
request: clonerequest@sanger.ac.uk
On Apr 30, 2001 this sequence version replaced gi:13396709.

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was either finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM: EMBL; SW: SWISSPROT; TR: TrEMBL; WP: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C-elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-504J9 is from the library RPI1-11.2 constructed by the group of Pleter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-504J9. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-504J9 is at 80148 in this sequence. The true right end of clone RP11-504J9 is at 100 in this sequence.


```
FEATURES
Source
Location/Qualifiers
1. 80247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-504J9"
/clone_lib="RPCT-11.2"
184. 1422
/note="L2 repeat: matches 875. .2419 of consensus"
repeat_region
1744. 1834
/note="MIR repeat: matches 166. .258 of consensus"
repeat_region
2950. 2977
/note="14 copies 2 mer ca 100% conserved"
repeat_region
3411. 3669
/note="LTR16C repeat: matches 93. .364 of consensus"
4372. 4704
/note="LTR1A1 repeat: matches 5. .365 of consensus"
misc_feature
complement(7821..8406)
/note="match: GSS: Em:A0550748"
8126. 8184
/note="MER5B repeat: matches 272. .332 of consensus"
repeat_region
8185. 8696
/note="LTR2D repeat: matches 1. .513 of consensus"
repeat_region
8779. 8817
/note="LTR2D repeat: matches 504. .542 of consensus"
repeat_region
8819. 8974
/note="MER58A repeat: matches 6. .162 of consensus"
misc_feature
complement(9443..9891)
/note="match: GSS: Em:B90633"
complement(9452..9883)
/note="match: GSS: Em:A0119695"
9782. 9894
/note="MIR repeat: matches 92. .205 of consensus"
repeat_region
10089. 10212
/note="MIR repeat: matches 135. .262 of consensus"
misc_feature
complement(10093..10671)
/note="match: GSS: Em:A0237061"
10223. 10282
/note="MIR repeat: matches 203. .262 of consensus"
misc_feature
complement(10315..10631)
/note="match: GSS: Em:A0200259"
10805. 11186
/note="match: GSS: Em:A007253"
11657. 11844
/note="L2 repeat: matches 2546. .2750 of consensus"
complement(11659..12084)
/note="match: GSS: Em:A0518361"
/note="match: GSS: Em:A0094678"
12083. 12588
/note="match: GSS: Em:A0691235"
12596. 12810
/note="MIR repeat: matches 23. .248 of consensus"
repeat_region
12920. 12986
/note="L1M45 repeat: matches 6228. .6295 of consensus"
repeat_region
12987. 13169
/note="L1M45 repeat: matches 5852. .6043 of consensus"
13168. 13280
/note="L1M48 repeat: matches 5710. .5832 of consensus"
13851. 17218
/note="L1P1 repeat: matches 2787. .6155 of consensus"
18903. 18953
/note="LTR2CB repeat: matches 439. .501 of consensus"
18984. 19433
/note="LTR2CB repeat: matches 1. .460 of consensus"
complement(19108..19690)
/note="match: GSS: Em:A0563935"
20683. 20801
/note="AluYo/FRAM repeat: matches 169. .288 of consensus"
20802. 20862
/note="AluJ/FLAM repeat: matches 3. .66 of consensus"
20912. 21033
/note="MIR repeat: matches 114. .237 of consensus"
repeat_region
21047. 21128
/note="41 copies 2 mer ta 63% conserved"
repeat_region
21676. 21859
/note="L2 repeat: matches 2130. .2208 of consensus"
21890. 23612
/note="L1P repeat: matches 841. .2570 of consensus"
23601. 26609
/note="L1P45 repeat: matches 3107. .6145 of consensus"
26620. 26789
/note="MER5A repeat: matches 1. .174 of consensus"
28521. 28627
/note="MER39 repeat: matches 5. .111 of consensus"
28656. 28837
/note="MER39b repeat: matches 54. .235 of consensus"
28838. 29142
/note="AluSc repeat: matches 4. .308 of consensus"
29143. 29494
/note="MER39b repeat: matches 235. .549 of consensus"
30529. 31344
/note="L1P43 repeat: matches 5331. .6146 of consensus"
32234. 32340
/note="L2 repeat: matches 2627. .2750 of consensus"
32399. 32578
/note="MIR repeat: matches 12. .203 of consensus"
32816. 32979
/note="MIR repeat: matches 91. .259 of consensus"
complement(33006..33450)
/note="match: GSS: Em:A0261227"
complement(33058..33444)
/note="match: GSS: Em:A0040453"
33312. 33596
/note="L1M4 repeat: matches 3473. .3773 of consensus"
34856. 34915
/note="30 copies 2 mer aa 68% conserved"
complement(34987..35547)
/note="match: GSS: Em:A0321598"
match: STS: Em:G55571"
complement(36488..36936)
/note="match: GSS: Em:B41073"
complement(36654..37090)
/note="match: GSS: Em:A0772584"
37102. 37591
/note="match: GSS: Em:A0489567"
complement(37581..38215)
/note="match: GSS: Em:A0418233"
37720. 37772
/note="L2 repeat: matches 2657. .2710 of consensus"
complement(37778..38215)
/note="match: GSS: Em:A0819773"
38218. 38732
/note="match: GSS: Em:A0599868"
38861. 39164
/note="AluSc repeat: matches 1. .305 of consensus"
39612. 39653
/note="21 copies 2 mer ac 90% conserved"
40854. 40909
/note="L2 repeat: matches 2655. .2710 of consensus"
42169. 42270
/note="MER5A repeat: matches 1. .111 of consensus"
42411. 42548
/note="MER5B repeat: matches 21. .181 of consensus"
42780. 42938
/note="MER5B repeat: matches 17. .177 of consensus"
43229. 43326
/note="MIR repeat: matches 40. .148 of consensus"
45383. 45428
/note="MER5B repeat: matches 18. .62 of consensus"
complement(47284..47971)
/note="match: STS: Em:G51995"
match: GSS: Em:A0269059"
complement(47537..48011)
/note="match: STS: Em:G36899"
complement(47723..48014)
misc_feature
```



```

|||||
Db 330 ACCCTTGACCTTACATCTGCGCCAGCAATGTTACATCCAGAGATCCATCCACAGC 389
Oy 121 GACCGTTCACTGACAAACAATCCCATGAGTTCTGATCCATGACGAGACATGCTGG 180
Db 390 GACCGTTCACTGACAAACAATCCCATGAGTTCTGATCCATGACGAGACATGCTGG 449
Oy 181 GATCTGACAAATTTACCTGACAGAGTAGCCACCTTCAGAAATGATTTGGATGAAT 240
Db 450 GATCTGACAAATTTACCTGACAGAGTAGCCACCTTCAGAAATGATTTGGATGAAT 509
Oy 241 TTGGAATCACAACCTACACTCCCAAAAGAGTAGGATGAGACTGTTTAACTGACT 300
Db 510 TTGGAATCACAACCTACACTCCCAAAAGAGTAGGATGAGACTGTTTAACTGACT 569
Oy 301 CTGACTTACCGCGCGTGTATCAGATATCAAGTCCCTTATGCTTTCAGAGTAGACACA 360
Db 570 CTGACTTACCGCGCGTGTATCAGATATCAAGTCCCTTATGCTTTCAGAGTAGACACA 629
Oy 361 AATCCCTGCTGTTTGAAGTGCACAGCAAGAGAAAATGGTGTGCTGGTTGTGAGTAA 420
Db 630 AATCCCTGCTGTTTGAAGTGCACAGCAAGAGAAAATGGTGTGCTGGTTGTGAGTAA 689
Oy 421 TGGAAACCTTGAGATGCCAGAGTCAAGTATTCATAGAGTAGCAAAAAGCATTTGAAT 480
Db 690 TGGAAACCTTGAGATGCCAGAGTCAAGTATTCATAGAGTAGCAAAAAGCATTTGAAT 749
Oy 481 CATACCTAGGCGCAACATTTTGAGAAATGTCATGATTAATAATTTGATTCCTACCAT 540
Db 750 CATACCTAGGCGCAACATTTTGAGAAATGTCATGATTAATAATTTGATTCCTACCAT 809
Oy 541 TCTGCTTGAATTTTATCTTCTCTTGAATAATTCATCCAGAGATTACATCAGGAA 600
Db 810 TCTGCTTGAATTTTATCTTCTCTTGAATAATTCATCCAGAGATTACATCAGGAA 869
Oy 601 AAGCTTAAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 870 AAGCTTAAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
Oy 661 AACTATGAGAAATTAATTCACAGAGATTCATTCATGATGAGAAATTAATTCACAG 720
Db 930 AACTATGAGAAATTAATTCACAGAGATTCATTCATGATGAGAAATTAATTCACAG 989
Oy 721 AGTACCTAGCAAAAGTATCTGAAGAGTGCACAAAACAAATTAATTAATTAATTA 780
Db 990 AGTACCTAGCAAAAGTATCTGAAGAGTGCACAAAACAAATTAATTAATTAATTA 1049
Oy 781 TTTAATGAGAGAGATTTCACTGTAATCTTCACAGATTTTGGAAATCAGATGAT 840
Db 1050 TTTAATGAGAGAGATTTCACTGTAATCTTCACAGATTTTGGAAATCAGATGAT 1109
Oy 841 TTGGCTTGGATCATGATGAGAAAGGATCAAGAAATTAATTAATTAATTAATTA 900
Db 1110 TTGGCTTGGATCATGATGAGAAAGGATCAAGAAATTAATTAATTAATTAATTA 1169
Oy 901 TGGTTTGGAAATTA 915
Db 1170 TGGTTTGGAAATTA 1184

```

```

RESULT 4
AF345993 1128 bp mRNA linear ROD 07-MAR-2001
LOCUS Rattus norvegicus alphas,3-fucosyltransferase IX (Fut9) mRNA,
DEFINITION complete cds.
ACCESSION AF345993
VERSION AF345993.1 GI:13242183
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

```

REFERENCE 1 (bases 1 to 1128)
AUTHORS Baboval,T., Henion,T., Kinnally,E. and Smith,F.I.
TITLE Molecular cloning of rat alphas,3-fucosyltransferase IX (Fuc-TIX)
and comparison of the expression of fuc-TIV and fuc-TIX genes
during rat postnatal cerebellum development
JOURNAL J. Neurosci. Res. 62 (2), 206-215 (2000)
PUBMED 11020213
REFERENCE 2 (bases 1 to 1128)
AUTHORS Smith,F.I. and Baboval,T.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) Biomedical Sciences, Eunice Kennedy Shriver
Center, 200 Trapelo Rd., Waltham, MA, USA
FEATURES
source
1..1128
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1128
/gene="Fut9"
49..1128
/codon_start=1
/product="alphas,3-fucosyltransferase IX"
/protein_id="AAK16591.1"
/db_xref="GI:13242184"
/translacion="MTSTKGLRPEFLIYCIILGCFVACLIYIRPTNSWVSPMESA
SVLMMKNEFSRKTDFYFNFTLLVWMPGQFDELSCQAMPNIGCHLTRSHLYNK
SHAVLIHHDDISWDLNLPQAPRPFQKIMWNLSPHTPQKSEIHELFNLTIRYR
DSDIQVPRGFLVTSVSPFVPEPSKELKIVSVNMPNPHAVKYEINLSKIEIHTY
GOAFGEIVDNKRLIPTISCKEFLSPENSINHDYITEKLYNAPFLAGSVVPLGPEREN
YENYIPADSEIHVEDFNSPELAKYLEKDKNNKLYLSYFNMRKDTVNLPRFESHA
CLACHVHRKHQKSVSNLEKFWFN"
BASE COUNT 324 a 256 c 228 g 320 t
ORIGIN
Query Match 86.6%; Score 792; DB 10; Length 1128;
Best Local Similarity 91.8%; Pred. No. 2,5e-198;
Matches 837; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Oy 4 AAACGATTAATTTAATGAACTACTATTCGCTGGGTGGCGCATTTGGCAGACC 63
Db 217 AAACGATTAATTTAATGAACTACTATTCGCTGGGTGGCGCATTTGGCAGACC 276
Oy 64 TTGACCTTACATCTGCCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAGGAC 123
Db 277 TTGACCTTACATCTGCCAAGCAATGTTCAACATCCAAAGATGCCATCTCACAGGAC 336
Oy 124 CGTTCACTGACAAATCCCATGCAATGCTGATCCATCACCAGACATCACTGGGAT 183
Db 337 CGCTCGTTGACAAATCCCATGCGTCTGATACACACAGACATCACTGGGAT 396
Oy 184 CCGCAAAATTTACCTGAGCAAGTACGACCCCTTCAGAAATGAAATTTGAGCAATTT 243
Db 397 CTGACTTAATTTACCTGAGCAAGTACGACCCCTTCAGAAATGAAATTTGAGCAATTT 456
Oy 244 GAATCACCAACTCACACTCCCAAAAGAGTGGCATGAGCACTGTTTAACTGACTG 303
Db 457 GAGTCAACCACTCACACCCCAAAAGAGTGGCATGAGCACTGTTTAACTGACTG 516
Oy 304 ACTTACCGCGGTATTCAGATATCCAGTGCCTTATGCTTCTTGAACGGTAAACCAAT 363
Db 517 ACTTATCGCGGTATTCAGATATCCAGTGCCTTATGCTTCTTGAACGGTAAACCAAG 576
Oy 364 CCCTGTGTTTGAAGTGCACCAAGAAATTTGTTGCTGCGGGGTTGAGTAATCTG 423
Db 577 CCCTGTGTTTGAAGTGCACCAAGAAATTTGTTGCTGCGGGGTTGAGTAATCTG 636
Oy 424 AACCTGAGCATGCCAGAGTCAAGTATTAATGAGTCAAGCAAAAGCATTTGAATCCAT 483
Db 637 AATCCGAGAGATGCCAGGATCAAGTATTAATGAGTCAAGCAAAAGCATTTGAATCCAC 696
Oy 484 AACTAGGCGCAACATTTGAGAAATGTCATGATTAATAATTTGATTCCTACCATATCT 543

```

Db	697	ACCTATGGGCAAGCATTGTGGAGAAATATGTGATGTATAAAAATCTGATTCGCCACATATCT	756
QY	544	GCTTGTAAATTTTATCTTTTCCATTTTGAAAAATTCATTCACAAGAGATTTACATCCAGCAAAAG	603
Db	757	ACTGTAAATTTTATCTTCTTCCTTTGAAAAACATTCACAAGAGATTTACATCCAGCAAAAG	816
QY	604	CTATTCAGTGCCTTTCTGSGTGGCTGCTGATCTGATCTGTGTTCTGGAGCAATCTAGGAAAAC	663
Db	817	CTCTACAAATGCATTTTGTGGCTGTCTGTACTGTTGTCTCTGGTCCGTCATCTTAGGAAAAAC	876
QY	664	TATGCAATTTATTTCCAGCAGATTCATTCATTCATGAGGAAATATTAACCTCCCACT	723
Db	877	TATGCAATTTATTTCCAGCTGATTCATTCATTCATGAGGAAATTTTAACCTCCCACT	936
QY	724	GAGTCAGCAAGTATCTGTAAAGAAAGTCGACAAAAACAATTAAGTATTAAGTCTGCTACTTT	783
Db	937	GAGTCGCAAAATATCTGTAAAGAGTTGACAAAAACAATTAAGTATTAAGTCTGCTACTTT	996
QY	784	AACGTGAGGAAGATTTTCACTGTAAATCTTCCACATTTTGGGAATTCACATGCAATGTTG	843
Db	997	AACGTGAGGAAGATTTTCACTGTAAACCTCCACGTTTGGGAATTCACATGCAATGCTTG	1056
QY	844	GCTTCCGATCATGTGAAAAAGGCATCAAGAAATTAAGTCTGTGTATTTAGAGAAATGG	903
Db	1057	GCATCGGATCATGTGAAAAAGGCATCAAGAAATTAAGTCTGTGTGCAATTTAGAGAAATGG	1116
QY	904	TTTTTGAATTTAA 915	
Db	1117	TTTTTGAATTTAA 1128	

	AB015426	2139 bp	mRNA	linear	ROD 23-JUN-1999
LOCUS	AB015426				
DEFINITION	Mus musculus Fut9 mRNA for alpha,3-fucosyltransferase IX, complete cds.				
ACCESSION	AB015426				
KEYWORDS	AB015426.1 GI:3702718				
SOURCE	Fut9; alpha,3-fucosyltransferase IX. Mus musculus cDNA to mRNA.				
ORGANISM	Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.				
REFERENCE	1 (sites)				
AUTHORS	Kudo,T., Ikebara,Y., Togayachi,A., Kaneko,M., Hiraga,T., Sasaki,K. and Narimatsu,H.				
TITLE	Expression cloning and characterization of a novel murine alpha,3-fucosyltransferase, mFuc-TX, that synthesizes the Lewis x (CD15) epitope in thymic and kidney U. Biol. Chem. 273 (41), 26729-26738 (1998)				
JOURNAL	2 (bases 1 to 2139)				
MEDLINE	98434586				
REFERENCE	Kudo,T. and Narimatsu,H.				
AUTHORS	Direct Submission				
TITLE	Submitted (09-JUN-1998) Takashi Kudo, Institute of Life Science, Soka University, Division of Cell Biology; 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan (E-mail:tkdud@et.soka.ac.jp, Tel:+81-426-91-2495, Fax:+81-426-91-9315)				
JOURNAL	Location/Qualifiers				
FEATURES	1..2139				
source	/organism="Mus musculus"				
gene	/db_xref="taxon:10090"				
	115..1194				
	/gene="Fut9"				
	115..1194				
CDS	gene="Fut9"				
	/codon_start=1				
	/product="alpha,3-fucosyltransferase IX"				
	/protein_id="AAA3352.1"				
	/db_xref="GI:3702719"				
	/translation="MTSKGILAPFLIVCIIGCFMACLIYRKPNNSVSPMSHSSVAHLIHRRSDLTNIPQARPPQKWMNNSEPLHPQKGIEHLFWLTYRRL"				

	DSDLDVQVEFLTYSINPFEVEVSREKLCYCWVSNMNEDEARAKYNELSKSIEHTY
	GAGFGEVNNDKNLIPRTSTCKFYLSASFKSLTHKKDITPERKLATNLAGSPVYLDSREN
	YENVIIPADSFDIHVEDFNFSPESELAKXLYKEVKNNKRLYSTIYNMKDKFVLNPRESHA
	CLACDHKRKHQEFKRSVGNLEKRFEN"
	2139
polyA_site	/note-"31 a nucleotides"
BASE COUNT	684 a 395 c 368 g 692 t
ORIGIN	
Query Match	86.5%; Score 791.8; DB 10;
Best Local Similarity ~	91.6%; Pred. No. 2,8e-198;
Matches 838; Conservative	0; Mismatches 77; Indels 0; Gaps 0;

QY	1	ACCAAAATGATATTTTAAATGAAGACATCTATTCGGTGGGTGGCCATTTGGGGAG	60
Db	280	ACAAAATGATATTTTAAACGAACATCTCGTTGGTATGGCCATTTGGGGAG	339
QY	61	ACCTTGACCTTAATCTCGCAAGCAATGTAAACATCCAAAGATGGCATCTCAACG	120
Db	340	ACCTTTGACCTTAATCTCGCAAGCAATGTCAATATCCAAAGGTGGCATCTCAACA	399
QY	121	GACCGTACGTATCAACAAATCCCATGCACTTCGTGATCCATCAACGAGATATAGTTGG	180
Db	400	GACCGCTATTTGACAAATATCCATCGGTGCTGTGATTCACCATAGAGATATAGCTGG	459
QY	181	GATCTGACAAATTTAAGTACAGAAAGTATGCGCACCCCTTCCAGAAATGGATTTGATGAT	240
Db	460	GATCTGATTAATCTACCTACAGAGCGCCGCCACCCCTTTCAGAAATGGATTTGATGAT	519
QY	241	TTGGAATCACCACACTCACACTCCCAAAAAGATGGCATTTGAGACCTGTTTAACTGACT	300
Db	520	TTAGAGTACCCACACTCACACCCCAAAAAGATGGCATTTGAAACCTGTCTCAACCTGACT	579
QY	301	CTGACTTACCCCGCTGATTCAGATATCCAACTGCTTATGAGCTTCTTGACGATTAAGACA	360
Db	580	CTAATCTTACCCCGGATTCAGATATCCAACTGCTTATGAGCTTCTTGACGATTAAGACA	639
QY	361	AATCCCTCGTGTGAAGTCCCAAGCAAAAGAAATGGTGTGCTGGTGTGAGTATAC	420
Db	640	AATCCCTTGTGTGAAGTCCCAAGCAAGGAAGTGTGTGCTGGTGTGAGTATAC	699
QY	421	TGGAACCTGAGCATGCGCAGATCAAGTATTCAAATGAGTATCAACAAAGCATGGAATC	480
Db	700	TGGAACCTGAGCATGCGCAGGATCAAGTATTCAAAGAGCTCAGCAAGATATTTGAATC	759
QY	481	CATACCTACGGGCAAGCATTTGGAGAAATATGTCATGATATATTTGATTTCCATACATA	540
Db	760	CACACCTATGSCCAAGCATTCGGAGAAAGCTGAAGATTAATAATCTGATTTCCACATA	819
QY	541	TCCTGCTTGAATTTATCTTTCCCTTTGAAAATTCATCCAAAGATTTACATCAGGAA	600
Db	820	TCCTACTTTGAATTTATCTTTCAATTTGAAAATTCATCCAAAGATTTACATCAGGAA	879
QY	601	AAGCATATCAATGCTTTTCTGGCTGGCTCTTACCTGTTGTTCTGGAGCATCTAGAGAA	660
Db	880	AAGCTCTCAATGCAATTTTGGCTGGTTCACTACCTGTTGTTCTGGAGCATCTAGAGAA	939
QY	661	AACATATGAAATTTATATCCAGAGATTCATTCATCTCTGGAAGATATATACCTCC	720
Db	940	AACATATGAAATTTATATCCAGTGAATTCATTCATCTGGAAGATTTATACCTCC	999
QY	721	AGTGAAGCAAAAGATCTGAGGAAGTGCAGAAAAAATTAAGTTATACCTTAGTTAC	780
Db	1000	AGTGAATGCAAAATATCTGAGGAAGTGCAGAAAAAATTAAGTTATACCTTAGTTAC	1059
QY	781	TTTAACTGAGAGAGATTTCACTGTAAATTTCTCAGATTTTGGAAATCAATGATGT	840
Db	1060	TTTAACTGAGAGAGATTTTAACTGTAAATCTTCCACGATTTTGGAAATCAATGATGT	1119
QY	841	TTGGCTTGCGATCATGTGAAGGATCAGAAATTAAGTCTGTTGGTATTTAGAGAA	900
Db	1120	CTGGATCTGCGACATGTGAAGGATCAGAAATTAAGTCTGTTGGTATTTAGAGAA	1179

QY 901 TGGTTTGGCAATTAA 915
DB 1180 TGGTTTGGCAATTAA 1194

RESULT 6
AB049819
LOCUS AB049819 2156 bp mRNA linear ROD 18-JAN-2002
DEFINITION Rattus norvegicus FUR9 mRNA for alpha1,3-fucosyltransferase IX,
complete cds.
ACCESSION AB049819
VERSION AB049819.1 GI:13591588
KEYWORDS
SOURCE Rattus norvegicus cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (sites)
REFERENCE 1 (sites)
AUTHORS Shimoda,Y., Tajima,Y., Osanai,T., Katsune,A., Kohara,M., Kudo,T.,
Narimatsu,H., Takashima,N., Ishii,Y., Nakamura,S., Osumi,N. and
Sanai,Y.
TITLE Pax6 Controls the Expression of Lewis x Epitope in the Embryonic
Forebrain by Regulating alpha 1,3-Fucosyltransferase IX Expression
JOURNAL J. Biol. Chem., 277 (3), 2033-2039 (2002)
PUBMED 11675393
REFERENCE 2 (bases 1 to 2156)
AUTHORS Sanai,Y.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-2000) Yutaka Sanai, Tokyo Metropolitan Institute
of Medical Science, Department of Biochemical Cell Research;
Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113-8613, Japan
(E-mail:sanaie@rinsoken.or.jp, Tel:81-3-3823-2101(ex.5233),
Fax:81-3-3828-6663)
FEATURES
source location/Qualifiers
1..2156
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
346..1425
/gene="FUR9"
346..1425
/gene="FUR9"
/product="alpha1,3-fucosyltransferase IX"
/protein_id="BAB40953.1"
/db_xref="GI:13591589"
/translation="MSTSKGILRPILYICIIIGCFVACLIYIKPTNSWSPWPSA
SVLKKNFESKTDYFNETITLWVWPGQFDLFCQAMNIGCHITTPRSILNK
SHAVLIHRDISMDLTNLDPQARPPQKIMNLSPTHPKSGIEHLFNLITRK
DSDIQVPGFELVSPVEFVSKSEKLVKMWNPENHAKVKNYELSKIEHTY
GQAFGEYVNDKMLIPTISCKEFLSEFENIHKDYITEKLYNAFLAGSVVVLGPSREN
YENIIPADSEIHVEDNSPELAKYLIKEDVKNKMLYLSYFNMRKDETVLPLRWESHA
CLACDHVKHQRKSVGNLEKMFNN"
BASE COUNT 605 a 470 c 437 g 644 t
ORIGIN

Query Match 86.4%; Score 790.4; DB 10; Length 2156;
Best Local Similarity 91.7%; Pred. No. 6.5e-198;
Matches 836; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 4 AAAAGTGATTTATTTAATGAAGTACTATTTCTGGTGGTGCCATTTGGCGAGACC 63
DB 514 AAAAGTGATTTATTTAATGAAGTACTATTTCTGGTGGTGCCATTTGGCGAGACC 573
QY 64 TTGACCTTACATCTGCCAGCAAGTGTTCATCATCCAAAGATGCATTCACAGAGGAC 123
DB 574 TTGACCTTACATCTGCCAGCAAGTGTTCATCATCCAAAGATGCATTCACAGAGGAC 633
QY 124 CGTTCACTGACAAACAATCCATGAGTCTGATCATCCAGGACATCAGTTGGGAT 183
DB 634 CGTCTGTTGACAAACAATCCATGAGTCTGATCATCCAGGACATCAGTTGGGAT 693
QY 184 CTGACAAATTTTACCTGACGAGCTAGGCCACCTTCACGAATGATTTGGATGAATTTTG 243

DB 694 CTGACTTAATCTTACCTGAGCAGGCGCCGCTTCAGAAAATGATTTGGATGAACCTTA 753
QY 244 GAATCACACACTCACACTCCCAAAAAGAGGATTTAGACATCTTTTAACCTGACTG 303
DB 754 GAGTCACACCACACTCACACCCCCAAAAGAGGATTTAGACATCTTTTAACCTGACTG 813
QY 304 ACTTACCGCGGTGATTCAGATATTCAGATGCCCTTATGCTTCTTGACGGTAAACCAAAAT 363
DB 814 ACTTATGCGCGGTGATTCAGATATTCAGATGCCCTTATGCTTCTTGACGGTAAACCAAAAG 873
QY 364 CCCTTGTGTTTAAAGTGGCCAAACCAAGAAATTTGGTGTCTGGGTTTGAGTAACTGG 423
DB 874 CCCTTGTGTTTAAAGTGGCCAAACCAAGAAATTTGGTGTCTGGGTTTGAGTAACTGG 933
QY 424 AACCTGAGCATCCAGAGCATAGTATTTACATGAGGTAAAGCAAAAGCATTTAAATCCAT 483
DB 934 AATCCGAGCATCCAGAGCATAGTATTTACATGAGGTAAAGCAAAAGCATTTAAATCCAT 993
QY 484 ACCTAGCGGCAACATTTGGAGATATGTCATGATPAAAAATTTGATTTCTACCATATCT 543
DB 994 ACCTAGCGGCAACATTTGGAGATATGTCATGATPAAAAATTTGATTTCTACCATATCT 1053
QY 544 GCTTTAAATTTTATCTTTCTTTGAAATTTCAATCCACAGATTTACATCAGGAAAG 603
DB 1054 ACTTGTAAATTTTATCTTTCTTTGAAATTTCAATCCACAGATTTACATCAGGAAAG 1113
QY 604 CTATPACATGCTTTCTGCGTGGCTGCTGTCATCGTGTGTTGGTGGACCATTTAGGAAAC 663
DB 1114 CTATPACATGCTTTCTGCGTGGCTGCTGTCATCGTGTGTTGGTGGACCATTTAGGAAAC 1173
QY 664 TATGAGATTAATATTCACAGATTCATTCATTCATGAGATTAATATTCATCCAGT 723
DB 1174 TATGAGATTAATATTCACAGATTCATTCATTCATGAGATTAATATTCATCCAGT 1233
QY 724 GACTAGCAAGATATCTGAAAGAGTCGACAAAACCAATATGATTTACTTACTTT 783
DB 1234 GACTAGCAAGATATCTGAAAGAGTCGACAAAACCAATATGATTTACTTACTTT 1293
QY 784 AACTGAGCAAGATATCTGAAAGAGTCGACAAAACCAATATGATTTACTTACTTT 843
DB 1294 AACTGAGCAAGATATCTGAAAGAGTCGACAAAACCAATATGATTTACTTACTTT 1353
QY 844 GCTTGGCATCATGTGAAGAGCATCAAGATATATAGTCTGTGATTTAGAGAAATGG 903
DB 1354 GCTTGGCATCATGTGAAGAGCATCAAGATATATAGTCTGTGATTTAGAGAAATGG 1413
QY 904 TTTTGGCAATTAA 915
DB 1414 TTTTGGCAATTAA 1425

RESULT 7
AF230460
LOCUS AF230460 1705 bp mRNA linear ROD 12-JUL-2000
DEFINITION Cricetusulus griseus alpha(1,3)fucosyltransferase type IX (Fut9)
mRNA, complete cds.
ACCESSION AF230460
VERSION AF230460.1 GI:9049663
KEYWORDS
SOURCE Chinese hamster.
ORGANISM Cricetusulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetusulus.
1 (bases 1 to 1705)
REFERENCE 1 (bases 1 to 1705)
AUTHORS Patnaik,S.K., Zhang,A., Shi,S. and Stanley,P.
TITLE Alpha(1,3)fucosyltransferases expressed by the gain-of-function
Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30
JOURNAL Arch. Biochem. Biophys. 375 (2), 322-332 (2000)
MEDLINE 20166953
PUBMED 10700388
REFERENCE 2 (bases 1 to 1705)

AUTHORS Patnaik,S.K.; Shi,S. and Stanley,P.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-2000) Cell Biology, Albert Einstein College of Medicine, Chanin 516, 1300 Morris Park Avenue, New York, NY 10461, USA

FEATURES

Source Location/Qualifiers
 1..1705
 /organism="Cricetulus griseus"
 /db_xref="taxon:10029"
 1..1705
 /gene="Fuc9"
 /note="cgFuc9"
 62..1141
 /gene="Fuc9"
 /note="Fuc-TIX"
 /codon_start=1
 /product="alpha(1,3)fucosyltransferase type IX"
 /protein_id="AAF82412.1"
 /db_xref="GI:9049664"
 /translation="MTSTSKGILRPLIVCVIILACFMACILITIKPTNSVSPMSA
 SSVLKMKRFEETKDYENETIIIVWMPGOTPDITSCAMENIOGCHLTPTDSLYRK
 SHAVLIHHRDISMDLTNPQARPPPOKIMMIESPTHPKSGIIEHLNITLTYRK
 DSDIOVPGRTLVSTNPVPEVPSKEKLVGVSNMPEHARVYIEHLSKSIETHT
 GOARGEVYNEKNLIPITSTCFYISFENSIRKDTITKLNALFAGSVYVGLGSRN
 YENIIPADSEIHVEDINSPSELAKILEVDNNKLYLSTFNMKDFVNLPRWESHA
 CLACDHRVHDEYSVGLKEMFN"

CDS
 62..1141
 /codon_start=1
 /product="alpha(1,3)fucosyltransferase type IX"
 /protein_id="AAF82412.1"
 /db_xref="GI:9049664"
 /translation="MTSTSKGILRPLIVCVIILACFMACILITIKPTNSVSPMSA
 SSVLKMKRFEETKDYENETIIIVWMPGOTPDITSCAMENIOGCHLTPTDSLYRK
 SHAVLIHHRDISMDLTNPQARPPPOKIMMIESPTHPKSGIIEHLNITLTYRK
 DSDIOVPGRTLVSTNPVPEVPSKEKLVGVSNMPEHARVYIEHLSKSIETHT
 GOARGEVYNEKNLIPITSTCFYISFENSIRKDTITKLNALFAGSVYVGLGSRN
 YENIIPADSEIHVEDINSPSELAKILEVDNNKLYLSTFNMKDFVNLPRWESHA
 CLACDHRVHDEYSVGLKEMFN"

BASE COUNT 515 a 336 c 320 g 534 t
 ORIGIN
 Query Match 85.7%; Score 783.8; DB 10; Length 1705;
 Best Local Similarity 91.0%; Pred. No. 3,6e-196;
 Matches 833; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

1 ACCAAACGATTTATTTAATGAACACTACTATCTGCTGTGGTGGCCATTGGCGAG 60
 227 ACAAAACGATTTATTTAATGAACACTATTCTTGGTGGTGGCCATTGGCGAG 286
 61 ACCCTTGACCTTACATCCGCGCAAGCAGATGTCACATCCAAAGATGCTTCCACAG 120
 287 ACCTTTGACCTTACATCCGCGCAAGCAGATGTCACATCCAAAGATGCTTCCACAG 346
 121 GACGCTGCTGTACACAAATCCCATGCTGATTCATCACCAGACATCACTGG 180
 347 GACGCTGCTGTACACAAATCCCATGCTGATTCATCACCAGACATCACTGG 406
 181 GATCTGACAAATTTACCTGACAGACCTGAGCCACCTTCCAGAAATGATTTGATGAAT 240
 407 GATCTGACAAATTTACCTGACAGACCTGAGCCACCTTCCAGAAATGATTTGATGAAT 466
 241 TTGAATCACCAACTCACACTCCCAAAAGATGTCATTTGACACTTTTAACCTGACT 300
 467 CTAGAGTCTCCACATCAACCCCAAAAGATGTCATTTGACACTTTTAACCTGACT 526
 301 CTGACTTACCGCGCTGATTCAGATATCCAGTGCCTTATGCTTCTTGACGGTAAACACA 360
 527 CTAAACATACCGCGCTGATTCAGATATCCAGTGCCTTATGCTTCTTGACGGTAAACACA 586
 361 AATCCCTGCTTGTGAAGTGCCAAAGAAATGCTGCTGCTGGTGTGAGTAAAC 420
 587 AATCCCTGCTTGTGAAGTGCCAAAGAAATGCTGCTGCTGGTGTGAGTAAAC 646
 421 TGGAAACCTGAGCATCCAGAGTCAAGTATTAACATGAGCAGCAAAAGATGGAATC 480
 647 TGGAAACCTGAGCATCCAGAGTCAAGTATTAACATGAGCAGCAAAAGATGGAATC 706
 481 CATACCTTACGGCAAGCATTTGGAGAAATATGTCAAATGATAAATTTGATTCCTACATA 540
 707 CATACCTTACGGCAAGCATTTGGAGAAATATGTCAAATGATAAATTTGATTCCTACATA 766
 541 TCTGCTTAAATTTATCTTCTTGAATAATTCATCAACAGATTTACATCAGGAA 600
 767 TCTACTTGAATTTATCTTCTTGAATAATTCATCAACAGATTTACATCAGGAA 826

QY 601 AAGCTATACAAATGCTTTTCTGGCTGGCTGTACCTGTTGTTCTGGGACCATCTAGGAA 660
 DB 827 AAGCTATACAAATGCTTTTCTGGCTGGCTGTACCTGTTGTTCTGGGACCATCTAGGAA 886
 QY 661 AACTATGAGAATTAATATCCAGAGATTCATTCATTCATGATGGAAGATTATTAACCTGCC 720
 DB 887 AACTATGAGAATTAATATCCAGAGATTCATTCATTCATGATGGAAGATTATTAACCTGCC 946
 QY 721 AGTAGCTTACCAAGATATCTGGAAGATGACAAACAAATTAATTAACCTAGTTAC 780
 DB 947 AGTAGCTTACCAAGATATCTGGAAGATGACAAACAAATTAATTAACCTAGTTAC 1006
 QY 781 TTAACTGAGAGAGATTTCTACTGTAATCTTCACAGATTTTGGGATACATGATGAT 840
 DB 1007 TTCAACTGAGAGAGATTTCTACTGTAATCTTCACAGATTTTGGGATACATGATGAT 1066
 QY 841 TTGGCTGCGATGATGAGAAAGCATCAAGATTAAGTCTGTTGATTTAGAGAA 900
 DB 1067 CTGGCTTGTATCACGTAATAAAGCACCACAGATTAAGTCTGTTGATTTAGAGAA 1126
 QY 901 TGGTTTGAATTTAA 915
 DB 1127 TGGTTTGAATTTAA 1141

RESULT 8
 AB035906 1707 bp DNA linear WRT 05-JAN-2002
 LOCUS
 DEFINITION Gallus gallus gene for cfu9, complete cds.
 ACCESSION AB035906
 VERSION AB035906.1 GI:18146865
 KEYWORDS cfu9.
 SOURCE Gallus gallus brain DNA.
 ORGANISM Gallus gallus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE
 1 (cites)
 Kaneo,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
 TITLE The evolutionary history of glycosyltransferase genes
 JOURNAL Unpublished
 2 (bases 1 to 1707)
 REFERENCE Kaneo,M., Saitou,N. and Kitano,T.
 AUTHORS Direct Submission
 JOURNAL Submitted (17-DEC-1999) Miki Kaneo, National Institute of
 Genetics, Laboratory of Evolutionary Genetics; Yata 1111, Mishima,
 Shizuoka 411-8540, Japan (E-mail:mikaneko@med.id.yamagata-u.ac.jp,
 Tel:81-559-81-6790, Fax:81-559-81-6789)
 COMMENT Sequence updated (29-Feb-2000).

FEATURES
 source Location/Qualifiers
 1..1707
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /tissue_type="brain"
 268..1347
 /gene="cfu9"
 268..1347
 /gene="cfu9-TIX"
 /codon_start=1
 /product="cfu9"
 /db_xref="GI:18146865"
 /translation="MTSTSKGILRPLIVCVIILACFMACILITIKPTNSVSPMSA
 SSVLKMKRFEETKDYENETIIIVWMPGOTPDITSCAMENIOGCHLTPTDSLYRK
 SHAVLIHHRDISMDLTNPQARPPPOKIMMIESPTHPKSGIIEHLNITLTYRK
 DSDIOVPGRTLVSTNPVPEVPSKEKLVGVSNMPEHARVYIEHLSKSIETHT
 GOARGEVYNEKNLIPITSTCFYISFENSIRKDTITKLNALFAGSVYVGLGSRN
 YENIIPADSEIHVEDINSPSELAKILEVDNNKLYLSTFNMKDFVNLPRWESHA
 CLACDHRVHDEYSVGLKEMFN"

BASE COUNT 517 a 338 c 329 g 521 t 2 others
 ORIGIN
 Query Match 72.8%; Score 666; DB 5; Length 1707;

Best Local Similarity 83.0%; Pred. No. 4.3e-165;
Matches 739; Conservative 0; Mismatches 155; Indels 0; gaps 0;

```

QY 2 CCAAACTGATTTATTAAGAACTACTATTCGTGTGGTGGCCATTTGGCAGA 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 434 CCAAACTGATTTATTAAGAACTACTATTTGGTTGGCTTTGGCCATTTGGCAGA 493
QY 62 CTTTGAACCTTACCTCCGCAAGCAATGTTCAACATCCAGAGATGCATCCACA 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 CATCGATCTACATCCCTGCCAAACGATGTTCACATCCATGAGGTGCATCTGACTAT 553
QY 122 ACCGTGACGTACCAAAATCCATGATCCATCCATCCATCCATCCATCCATCC 181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 ACCGTGACGTATTAACAAATCCATCCATCCATCCATCCATCCATCCATCCATCC 613
QY 182 ATCTGACAAATTTACCTCAGCAAGCTGACCCCTTCAGAAATGGATTTGGATTT 241
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 ATCTGATTAATTTACCTCAGCAAGCTGACCCCTTCAGAAATGGATTTGGATTT 673
QY 242 TGGATACCAACTCAGCTCCCAAAAAGATGGCATTTGACACTTTTAACTGACTC 301
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 TGGATATCCAACTCAGCTCCCAAAAAGATGGCATTTGACACTTTTAACTGACTC 733
QY 302 TGAATACCCGCGTGAATTCAGATTCAGATGCTTATGGCTTTGACGCTAAGCACA 361
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 TGAATACCCGCGTGAATTCAGATTCAGATGCTTATGGCTTATGATAGTTGCTACG 793
QY 362 ATCCCTTCGTTGTTGAAGTCCCAAGAGAAATGCTGCTGGTGGTGTGATTAAT 421
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 GCGCTTTACATATGAAGTGCAGATTAAGAAATTTGGTTTGGTTGTAAGTAACT 853
QY 422 GGAACCTGAGCATGCAAGTCAAGATTAATCAATGAGCTAAGCAAAAGCATTAATCC 481
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 GGAACCTGAGCATGCGTCAAGATTAATCAATGAGCTAAGCAAAAGCATTAATCC 913
QY 482 ATACCTGAGGCAAGATTTGGAGATATGCAATGATTAATTTGATTTCTTCCATAT 541
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 ATACCTGAGGCAAGCTTTGGAGATATGCAATGATTAATTTGATTTCTTCCATAT 973
QY 542 CTGCTGTAATTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 601
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 CCAATTCGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1033
QY 602 AGCTAACAATGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 AACTTCAATGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
QY 662 ACTATGAGATTTATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTC 721
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 ACTATGAGATTTATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTC 1153
QY 722 GTGAGTACGAAGATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTC 781
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 GAGAGCTTGCAGATATCTTCTGATGCTTGCAGCAAAATTAACAGATGATCTTACT 1213
QY 782 TTAACGTGAGAGATTTCACTGTAATCTTCCAGATTTTGGGAATTCACATGCTAT 841
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 TCAACTGAGAGAGATTTCTCACTGATCTTCTTAAATCTTGGGAATTCATGCTAT 1273
QY 842 TGGCTTCGATTCATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTC 901
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1274 TGGCTTCGATTCATTCAGCAAGATTCATTCATTCATTCATTCATTCATTCATTC 1333
QY 902 GGTTTTGAATTA 915
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1334 GGTTTTGAATTA 1347

```

RESULT 9
AB035905
LOCUS
DEFINITION Xenopus laevis gene for xFUT9, complete cds.
ACCESSION AB035905

1080 bp DNA linear VRT 05-JAN-2002

VERSION AB035905.1 GI:18146863
KEYWORDS xFUT9.
SOURCE Xenopus laevis DNA.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae; Xenopus.

REFERENCE
AUTHORS Kaneko,M., Nishihara,S., Kitano,T., Narimatsu,H. and Saitou,N.
TITLE The evolutionary history of glycosyltransferase genes
JOURNAL unpublished
REFERENCE 2 (bases 1 to 1080)
AUTHORS Kaneko,M., Saitou,N. and Kitano,T.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1999) Miki Kaneko, National Institute of Genetics, Laboratory of Evolutionary Genetics, Yata 1111, Mishima, Shizuoka 411-8540, Japan (E-mail: mkaneko@med.nid.yamagata-u.ac.jp, Tel:81-559-81-6790, Fax:81-559-81-6789)
COMMENT Sequence updated (29-Feb-2000).
FEATURES
source
location/Qualifiers
1..1080
/organism="Xenopus laevis"
/db_xref="taxon:8355"
gene
1..1080
/gene="xFuc-TIX"
CDs
1..1080
/gene="xFuc-TIX"
/product="xFUT9"
/protein_id="BAB82488.1"
/db_xref="GI:18146864"
/translation="MTSASKGLIRAFVAVSLVLCIMCLLYKPKPNMWSPIESA
HSVLKKNFESKONYNEVTVLIVWPGQPELESCOTLENYGCHLTDBALYK
SHAVLJHNDLISMDLNLPOGLRPPROKPMWMLSPPTPKSGIEH.FNTLTYSR
DSDIQVPGVMVSTKPKPDEVSKEKPCWYVSNMPEHARKYINELIKETITTY
GAFGEYLSDKSLPTIASCKFYLSNTHKDYITKLYNALASVIVLSPREN
YENVIPADSFTHVEDFLSPRELDHLMINKDTEOYLSTYNNRKHTVNMPEWESHA
CLACDVHVKROEKFVSNLEKWMN"

BASE COUNT 332 a 231 c 202 g 315 t
ORIGIN

Query Match 63.5%; Score 581.2; DB 5; Length 1080;
Best Local Similarity 77.2%; Pred. No. 1e-14;
Matches 706; Conservative 0; Mismatches 208; Indels 0; gaps 0;

```

QY 2 CCAAACTGATTTATTAAGAACTACTATTCGTGTGGTGGCCATTTGGCAGA 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 167 CTAAACGAATTTATTAAGAACTCGTCTTAATTTGGTGTGGCTTTGGCAGA 226
QY 62 CTTTGAACCTTACCTCCGCAAGCAATGTTCAACATCCAGAGATGCATCCACA 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 CTTTGAACGAATTCCTCGCAAGCATTTGTTAATTTATTTATTTATTTATTTATTT 286
QY 122 ACCGTGACGTGACAAATCCATGATTCGATTCATTCACCGAGACATCAGTTGG 181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 ACCGCGCACTCTACCAAGATTCATCTCTCTCTTCTTCCACAGAGACATTAAGCT 346
QY 182 ATCTGACAAATTTACCTCAGCAAGCTGAGCCACCTTCAGAAATGGATTTGGATTA 241
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 ATCTCATTAACCTGCCCCCAACAGCTGAGGCCACTTTCAGAAATGGATTTGGATTA 406
QY 242 TGGATACCAACTCAGCTCCCAAAAAGATGGCATTTGAGCACTTTTAACTGACTC 301
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 TGGATACCAACACATACGCACAAAAAGATGGCATTTGAGCACTTTTAACTGACTC 466
QY 302 TGAATACCCGCGTGAATTCAGATTCAGATGCTTATGGCTTTGACGCTAAGCACA 361
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 TGAATACCCGCGTGAATTCAGATTCAGATGCTTATGGCTTTGACGCTAAGCACA 526
QY 362 ATCCCTTCGTTGTTGAAGTGCAGCAAGAAATGCTGCTGGTGGTGTGATTAAT 421
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 527 AACCATTTGACTTTGAGAGTGCAGCAAGCAAGATTTGCTGTGGTGTGATTAAGT 586

```



```

OY 422 GGAACCCGAGCAGCCAGCAGCAGTATATACATGAGCTAAGCAAAAGCATTGAATTC 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 587 GGAACCCGAGCAGCAGCAGCAGCAGTATATACATGAGCTAAGCAATGAATTA 646
OY 482 ATACCTACGGCAGCAGCAGCAGTATATGTCATGATATATATATATATATATAT 541
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 647 TAACTTATGGGAGCAGCAGCAGTATGTCATGAGCAGCAGCAGCAGCAGCAGTAT 706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 542 CTGCTTGAATATATATATATATATATATATATATATATATATATATATATATAT 601
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 707 CATCTTGAATATATATATATATATATATATATATATATATATATATATATATAT 766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 602 AGCTTATACATATATATATATATATATATATATATATATATATATATATATAT 661
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 767 AACTTATATATATATATATATATATATATATATATATATATATATATATATAT 826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 662 ACTATGAGATATATATATATATATATATATATATATATATATATATATATATAT 721
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 827 ACTATGAGATATATATATATATATATATATATATATATATATATATATATATAT 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 722 GTGAGCTAGCAAGATATATATATATATATATATATATATATATATATATATAT 781
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 887 GAGATATATATATATATATATATATATATATATATATATATATATATATATAT 946
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 782 TTAATGAGAGAGATATATATATATATATATATATATATATATATATATATATAT 841
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 947 ACAACTGAGAGAGATATATATATATATATATATATATATATATATATATATATAT 1006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 842 TGGCTTGGATCATATATATATATATATATATATATATATATATATATATATATAT 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1007 TAGCTTGGATCATATATATATATATATATATATATATATATATATATATATATAT 1066
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 902 GGTATTGGAATTA 915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1067 GGTCTGCACTAA 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AC100674 71396 bp DNA linear HTG 22-NOV-2001
AC100674 Mus musculus clone RP23-167K24, LOW-PASS SEQUENCE SAMPLING.
AC100674.1 GI:17048040
HTG: HTGS.PHASE0.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 71396)
Birten,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-167K24
Unpublished
2 (bases 1 to 71396)
Birten,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukigalter,B.,
Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,J.,
Choepey,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,I., Grand-pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labrecque,K.,
Lamazeres,R., Landers,T., Lenockky,J., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Meneu,C., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nobu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,N., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
```

TITLE
JOURNAL
COMMENT

```

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/xw/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L15932
Center Clone name: 167_K24
-----
* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
789 888: gap of 100 bp in length
889 1702: contig of 814 bp in length
1703 1802: gap of 100 bp
1803 2642: contig of 840 bp in length
2643 3637: gap of 100 bp
3638 3737: contig of 895 bp in length
3738 4635: gap of 100 bp
4636 4735: contig of 898 bp in length
4736 5657: gap of 100 bp
5658 5757: contig of 922 bp in length
5758 6652: gap of 100 bp
6653 6792: gap of 100 bp
6793 7703: contig of 935 bp in length
7704 7803: gap of 100 bp
7804 8742: contig of 939 bp in length
8743 8842: gap of 100 bp
8843 9787: contig of 945 bp in length
9788 9887: gap of 100 bp
9888 10819: contig of 932 bp in length
10820 10919: gap of 100 bp
10920 11808: contig of 889 bp in length
11809 11908: gap of 100 bp
11909 12830: contig of 922 bp in length
12831 12930: gap of 100 bp
12932 13847: contig of 917 bp in length
13848 13947: gap of 100 bp
13948 14886: contig of 938 bp in length
14887 14986: gap of 100 bp
14987 15911: contig of 925 bp in length
15912 16011: gap of 100 bp
16012 16903: contig of 892 bp in length
16904 17003: gap of 100 bp
17004 17906: contig of 903 bp in length
17907 18006: gap of 100 bp
18007 18931: contig of 925 bp in length
18932 19031: gap of 100 bp
19032 19945: contig of 914 bp in length
19946 20045: gap of 100 bp
20046 20951: contig of 906 bp in length
20952 21051: gap of 100 bp
21052 21949: contig of 898 bp in length
21950 22049: gap of 100 bp
22050 22950: contig of 901 bp in length
```


QY	718	CCCGATGAGCGAAGGCTGTAAGGAGGAGTGCACAAAACATTAAGTT-ATACCTTAG	776
Db	66261	CCCCGATGAGTGAACAATATCTCGAAGAGTGCACAAAACATTAAGTTAGTCTTAG	66320
QY	777	TTACTTTAACGCGAGGAGGAGGATTTTC	801
Db	66321	TTACTCTACTCGAGGAGGAGGATTC	66345
RESULT 11			
LOCUS	AB023627		
DEFINITION	AB023627	1567 bp	DNA
ACCESSION	AB023627		
VERSION	AB023627.1	GI:4587295	
KEYWORDS	alpha(1,3)fucosyltransferase.		
SOURCE	Danio rerio DNA.		
ORGANISM	Danio rerio		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
AUTHORS	Kageyama, N., Natsuka, S. and Hase, S.		
TITLE	1 (sites)		
JOURNAL	Molecular cloning and characterization of two zebrafish embryogenesis		
REFERENCE	J. Biochem. (1999) In press		
AUTHORS	2 (bases 1 to 1567)		
TITLE	Kageyama, N., Natsuka, S. and Hase, S.		
JOURNAL	Submitted (12-FEB-1999) Shinji Natsuka, Osaka University, Graduate School of Science, Machikaneyama 1-1, Toyonaka, Osaka 560-0038, Japan (E-mail: natsuka@chem.sci.osaka-u.ac.jp, Tel: 81-6-6850-5381, Fax: 81-6-6850-5383)		
FEATURES			
source	Location/Qualifiers		
gene	1. 1567		
CDS	/organism="Danio rerio"		
	/db_xref="taxon:7955"		
	177..1352		
	/gene="zfT1"		
	177..1352		
	/gene="zfT1"		
	/codon_start=1		
	/product="alpha(1,3)fucosyltransferase"		
	/protein_id="BA06706.1"		
	/db_xref="GI:4587295"		
	/translation="MDKRTLPSPKAAQVYIITMLSPVCFEYVNPNTPEKFP		
	AVDHTNCSTECLELDVFKRQNTKTKASNDSPPTAPNPAPAKARDDETLIA		
	WYDFGARFDLGVGSGOFNIHGCHLIDRSMIOKAGVHHHRDLSADIPQPRRLD		
	KIWMNKESSPSHNPNI.LINDGNLISFSPKSD.IPVYGRILDAIDPDKNTTFRD		
	KIVQWIVSNFQEHKRSQYVNELVKHKIKKAGCHGNHNDDDKGVNYSKRTAKD		
	ENSLDHNRYPTFKELFNPLALGVTAIPVSPVSGSDNVEEIPDPAITLHDDPPTKELADHL		
	KSLDNEEDQRYOYRVMKRFHIFMSSEFGEHACRICDYLKRKKHVIATIDLKGMVFWG-		
BASE COUNT	499 a	293 g	441 t
ORIGIN	334 c		
Query Match	30.4%	Score 277.8;	DB 5;
Best Local Similarity	58.5%;	Pred. No. 1.2e-62;	
Matches 527;	Conservative	0;	Mismatches 362; Indels 12; Gaps 2;
QY	18	TAAATGAACACTACTATTCGTGTGTGGTGGTGCACATTTGGGCAACACCTTGACCTTACATC	77
Db	461	TCAAGAAACCTTCCTGTGTGATCTGGGTGTATCCCTTGGGACGACGCTTTGACTTGGGAGT	520
QY	78	CTGCACAGCATGTTTCAACATCCAGAGATGCATTCACAAAGGACCGCTTGACTCTAGTACA	137
Db	521	GTGTGGTTCGATGATTCACACCGCTGTCTTAAACAGACGACGAGAAATATATATCA	580
QY	138	CAATTCACATGCAGTTTGTGATTCATCACCCAGACATCATGTTGGGATCTGACAAATTTACC	197
Db	581	GAAGCGCATGAGATGTATGTTTCAACACGAGATCTGAGCGCGG-----ATTGGC	631

QY	Db	712	GAGAACTCAGAGCAACAAGACTACATCAGTGAAGAAGCTCTGGAGGAATGCCCTGACCT	771
QY	625	GGCTCTGACTCTGTTGTTCTTGAGACCATCTGAGGAAAACTATGAGATTAATATTCACGA	684	
Db	772	GGCACACATCCCAAGTGTCTGGGGCCACCTCGAGAAAACTATGAGCACTCTTGCCCT	831	
QY	685	GATTCATTCATTCATGTCGAGATTAATTAACCTCCCATGAGCGACGAACTATCTGAG	744	
Db	832	GACTCTTCATTCACAGTGTGATGATCTTGCCAGTGTGATGAACTGGCGCAGTACTCTGAG	891	
QY	745	GAACTCGACAAAACCAATATGATTAATACCTTACTTAACTTGAGAGAG	795	
Db	892	GAAGTACAGACGAGACACCGACAGATGACGAGGCTACTTCCAGTGGCGAGG	942	
RESULT 15				
GGU73678		2657 bp	DNA	linear
LOCUS				VRT 20-DEC-1996
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
TITLE				
REFERENCE				
AUTHORS				
JOURNAL				
MEDLINE				
AUTHORS				
JOURNAL				
FEATURES				
SOURCE				
gene				
CDS				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches				
QY	28	ACTATTCGTGTGTGGGTGTGGCATTTGGGCGAGACCTTTGACCTTACATCTCTCCACGA	87	
Db	341	ACGGTGTGTGTGTGGGAGCCCTTCGCGCCGCCCTGGCGCGCGACGACTGGCGAGG	400	
QY	88	ATGTTCAACATTCGAAGATGCCATTCACACACGACCGTTCACCTTACATCAACAATCCAT	147	

Db	401	CGATCAAAATCAACGGGCTCCCTCTCGAGCGCCGACCGGGGGCGGTACGGGAGGCTCGG	460
Oy	148	GCAGTCTGATTCATCAACGAGACATCACTAGTTGGATCTGACAAATTAACCTCAGCAAG--	205
Db	461	GCGGTGCTCTCCACACCGCAGACTGGCGCTGCACGCGCGCAGGGGCTGCCCCGGGG	520
Oy	206	-----CTAGGCCACCCCTTCCAGAAATGGATTTGGATGAATTTGGAAATCACCACTCAC	258
Db	521	CCCCCGCGGACCCCGCGCGAGCGGTGGTGGATMACTTGAGATCGAGCCCTCGGCAC	580
Oy	259	ACTCCGCAAAAGATGGCATTTGAGCACTTGTTAACCTCAGCTGACTTACCGCGCTCAT	318
Db	561	TCCCCCGGGGCTCGGGGCGTGGCGGGCTCTTCACTGAGACATGTGCTGACGAGGGAC	640
Oy	319	TCAGATATCCAAAGGCGCTTATGAGCTTCTTGACGGTAAAGCAAAATCCCTTGGTTGTAA	378
Db	641	TCGAGACGTTCTTGCGCCCTACGGGTACTCTTACGACCGCGCGTCCGCGGGCCCTTCTGTG	700
Oy	379	GTCGCACGCAAAAGAAATTTGCTGTGCTGGGTGTGAGTACTGCAACCTGACGATCC	438
Db	701	CTGGCTGCAAAACCGGGCGTGGTGGCTGGGTATAGCAACTGGAGACGAGAGCAGCC	760
Oy	439	AGAGCTCAAGATTACAAATGACTAAGCAAAAAGCATTTGAATTCATACCTACGGGCAACA	498
Db	761	CGCGGCGCTACTACGGGACGTGAAGAGACACTGCCCATCGAGTGCAGG--CGG	817
Oy	499	TTTGGAGAAATATGTCAATGATTAATAATTGATTCCTACCATCTGCTTGTAAATTAT	558
Db	818	CGCGGGATGGCGCTGCTGAGGGGACAGGTGCTAMACGGTGTGCGGCTTACAAAGTTCTAC	877
Oy	559	CTTTCTTTGAAATTTCAATCCAAAGAGATTACATACGAGAAAGCTAT---ACAATCT	615
Db	878	CTGGCTTTCGAAATCTCCAGACACGAGATACATACGAGAAAGCTGTGGAAAGACCC	937
Oy	616	TTTCTGGCTGCTGTACCTGTGTGTCTGCGACATCTAGGGAATAACATAGAAATAT	675
Db	938	TTTCGCGCGACGGGGTGGCTGTGTGTCTGTGCGCCCCGAGGGCCAACTACGAGCGCTTC	997
Oy	676	ATTCAGACAGATTCATTTCATTTCATGTGGAAGATTATACCTGCCAGTGAAGCTAGCAAG	735
Db	998	ATCCCGCGCGACTCTTCATCAACGTGAGACTTCGCCAGCGCAAGGTGTGTGGCCAC	1057
Oy	736	TATCTGAAGAGATCGCAAAAAACAATAGTATATACCTTAGTTACTTTACTGAGAGAG	795
Db	1058	TACTCTGAAGTTCTCTCGATTAATAAACGCCAGCTACAGAGAGGTATTTGCTCGCGGAGAC	1117
Oy	796	GATTTCACTGTAATCTTCCACGATTTTGGGAATTCATGATGATTTTGGCTTGCATAT	855
Db	1118	AAGATGAGAGTCCAGTCAAGCTTCTCTGGATGAGCACTTACTTCGCAAGGTTTGCACGCC	1177
Oy	856	GTGAAAGG 864	
Db	1178	GTGAGAGCG 1186	

Search completed: October 6, 2002, 04:44:21
Job time : 1255.59 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:38:25 ; Search time 102.682 Seconds
(without alignments)
15299.444 Million cell updates/sec

Title: US-09-744-748-5_COPY_460_1374

Perfect score: 915
Sequence: 1 accaaacacgattatttaa.....agaatggtttggaattaa 915

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/hold-geneseq/geneq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	915	100.0	2676	21	AAZ92647 Human alpha-1,3-fu
2	915	100.0	2832	21	AAZ92646 Human alpha-1,3-fu
3	791.8	86.5	2036	21	AAZ92672 Murine alpha-1,3-f
4	791.8	86.5	2170	21	AAZ92645 Murine alpha-1,3-f
5	127.4	13.9	1814	18	AAZ95050 Murine myeloid-lin
6	127	13.9	3594	18	AAZ85219 Mouse alpha-fucosy
7	114.4	12.5	1701	15	AAZ77732 Human alpha-1,3-fu
8	114.4	12.5	1701	21	AAZ21136 Human low adenosin
9	114.4	12.5	1701	21	AAZ35014 Human adenosine re

10	114.4	12.5	6944	21	AAZ1137 Human low adenosin
11	114.4	12.5	6944	21	AAZ35015 Human adenosine re
12	113.4	12.4	1654	15	AAZ56905 pCDNA1-alpha-(1-3)
13	113.4	12.4	1654	18	AAZ61680 Human alpha(1,3)-f
14	112.2	12.3	1086	15	AAZ69912 Lewis blood group
15	112.2	12.3	2043	18	AAZ80111 Fucosyltransferase
16	112.2	12.3	2043	12	AAZ33330 Glycosyltransferase
17	112.2	12.3	2043	15	AAZ6906 DNA encoding a gly
18	112.2	12.3	2043	18	AAZ76769 Human alpha 1,3/4
19	112.2	12.3	2043	18	AAZ61675 Human alpha(1,3/1,
20	111.8	12.2	1126	21	AAZ21135 Human low adenosin
21	111.8	12.2	1126	21	AAZ35013 Human adenosine re
22	111.8	12.2	1316	15	AAZ56911 DNA encoding a gly
23	111.8	12.2	1316	18	AAZ61679 Human alpha(1,3)-f
24	103.6	11.3	795	17	AAZ13798 CDX, a MIRA Invol
25	102.4	11.2	2175	11	AAZ06691 Human low adenosin
26	100.8	11.0	1256	21	AAZ21134 Human adenosine re
27	100.8	11.0	1256	21	AAZ35012 Human adenosine re
28	100.8	11.0	1400	13	AAZ31436 Encodes a Hela cel
29	100.8	11.0	1488	15	AAZ56910 DNA encoding a gly
30	100.8	11.0	2175	12	AAZ14382 Human CDNA clone 7
31	100.8	11.0	2175	24	AAZ17082 Clone 1 encoding 1
32	100.8	11.0	2861	12	AAZ14383 Human low adenosin
33	100.8	11.0	2861	21	AAZ21133 Human adenosine re
34	100.8	11.0	2861	21	AAZ35011 Human adenosine re
35	100.8	11.0	2861	24	AAZ17083 Human CDNA clone 1
36	100.8	11.0	3647	15	AAZ56909 GDP-Fuc:beta-D-Gal
37	100.8	11.0	3647	18	AAZ61678 DNA encoding a gly
38	100.8	11.0	3647	15	AAZ56909 Human alpha(1,3)-f
39	97.6	10.7	2134	18	AAZ59506 Human myeloid deri
40	66.6	7.3	1378	23	ABL11867 Drosophila melanog
41	66.6	7.3	1429	23	ABL11866 Drosophila melanog
42	52.2	5.7	1353	21	AAZ51685 C. elegans alpha-1
43	44.4	4.9	676	22	AAZ1344 CDNA encoding nove
44	44.4	4.9	1479	24	ABA05334 Human fucosyltrans
45	44.4	4.9	2557	24	ABA05333 Human fucosyltrans

ALIGNMENTS

RESULT 1	AAZ92647	standard; cDNA; 2676 BP.
ID	AAZ92647	
XX	AAZ92647	
AC	AAZ92647	
XX	05-JUN-2000	(first entry)
DE	Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO:5.	
XX	Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;	
KW	brain; kidney; recombinant expression; transgenic animal; knockout	
KW	animal; FUC-TIV; drug screening; inhibitor; potentiator; diagnosis;	
KW	treatment; cancer; human; ss.	
XX	Homo sapiens.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	CDS	295..1374
FT		/*tag= a
FT		/product= "Human alpha-1,3-fucosyltransferase"
PN	WO200006708-A1.	
XX		
PD	10-FEB-2000.	
XX		
PF	29-JUL-1999;	99WO-JP04092.
XX		
PR	29-JUL-1998;	98JP-0213823.
XX		
PA	(RYOW) KYOWA HAKKO KOGYO KK.	
XX		

PI Narimatsu H, Kudo T, Sasaki K;
 XX WPI: 2000-183120/16.
 DR P-PSDB: AAY80996.
 XX
 XX Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x
 PT and kidney disease and cancer. -
 PT epitoide and an antibody recognizing it useful for diagnosis of brain
 PT
 XX
 XX Claim 4; Page 143-150; 172pp; Japanese.

XX The invention relates to a novel alpha-1,3-fucosyltransferase which
 CC transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglucosamine
 CC (generating the Lewis x or y epitope). It does not transfer a fucose
 CC moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglucosamine
 CC and therefore does not generate the sialyl-Lewis x epitope. The
 CC invention also relates to DNA sequences encoding alpha-1,3-
 CC fucosyltransferase and expression vectors and host cells comprising
 CC these DNA sequences. The invention additionally encompasses the
 CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC FUC-TV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequences
 CC AA92646 and AA92647 represent cDNAs encoding human
 CC alpha-1,3-fucosyltransferase (AAY80996).
 XX
 XX Sequence 2676 BP; 819 A; 527 C; 476 G; 854 T; 0 other;

Query Match 100.0%; Score 915; DB 21; Length 2676;
 Best Local Similarity 100.0%; Pred. NO. 7.8e-255;
 Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAAGCTGATTTATTAATGAAGTACTGATGCTGGTGGGCGCATTTGGGCGAG 60
 DB 460 ACCAAAGCTGATTTATTAATGAAGTACTGATGCTGGTGGGCGCATTTGGGCGAG 519
 OY 61 ACCTTTGACCTTACATCCCTGCCAAGCAATGTTCAACATCCCAAGGATCCATTCACAAAG 120
 DB 520 ACCTTTGACCTTACATCCCTGCCAAGCAATGTTCAACATCCCAAGGATCCATTCACAAAG 579
 OY 121 GACCGTTCAGTACCAACAAATCCCATGAGTTGATTCATCCAGGACATCAGTTGG 180
 DB 580 GACCGTTCAGTACCAACAAATCCCATGAGTTGATTCATCCAGGACATCAGTTGG 639
 OY 181 GATCGACAATTTACCTCAGCAAGTGGGCGACCTTCCAGGAATGATTTGGATGAAT 240
 DB 640 GATCGACAATTTACCTCAGCAAGTGGGCGACCTTCCAGGAATGATTTGGATGAAT 699
 OY 241 TTGGAATGACCAACACTCCCAAGGATGGGCAATGAGCACTGTTAACTGACACT 300
 DB 700 TTGGAATGACCAACACTCCCAAGGATGGGCAATGAGCACTGTTAACTGACACT 759
 OY 301 CTGACTTACCGCGGTGATTCAGATATCCAGTGCCTTATGGCTTCTGACGGTAAGCA 360
 DB 760 CTGACTTACCGCGGTGATTCAGATATCCAGTGCCTTATGGCTTCTGACGGTAAGCA 819
 OY 361 AATCCCTTCGTGTTGAAGTCCCAAGCAAGAAATGTTGCTGCTGGTGTGAGTAAC 420
 DB 820 AATCCCTTCGTGTTGAAGTCCCAAGCAAGAAATGTTGCTGCTGGTGTGAGTAAC 879
 OY 421 TGAACCTTGAGCATGCGAGAGTCAAGTATTAATGAGTAAAGCAAAAGCATTTGAATC 480

DB 880 TGAACCTTGAGCATGCGAGAGTCAAGTATTAATGAGTAAAGCAAAAGCATTTGAATC 939
 OY 481 CATACCTACGGGCAAGCATTTGGAGAAATATGTAATGATAAAATTTGATTCCTACCAT 540
 DB 940 CATACCTACGGGCAAGCATTTGGAGAAATATGTAATGATAAAATTTGATTCCTACCAT 999
 OY 541 TCTGCTGTAAATTTTATCTTCTCTTGAATTAATTCATCCAAAGATTTACATCAGGAA 600
 DB 1000 TCTGCTGTAAATTTTATCTTCTCTTGAATTAATTCATCCAAAGATTTACATCAGGAA 1059
 OY 601 AAGCTATTCATATGCTTTTCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 1060 AAGCTATTCATATGCTTTTCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119
 OY 661 AACTATGGAATTTATTTCCAGGATTCATTCATTCATTCATTCATTCATTCATTCATTC 720
 DB 1120 AACTATGGAATTTATTTCCAGGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1179
 OY 721 AACTATGGAATTTATTTCCAGGATTCATTCATTCATTCATTCATTCATTCATTCATTC 780
 DB 1180 AACTATGGAATTTATTTCCAGGATTCATTCATTCATTCATTCATTCATTCATTCATTC 1239
 OY 781 TTTAACTGGAGAGAGATTTCTGTAATCTTCCAGATTTTGGAAATCAGATGATG 840
 DB 1240 TTTAACTGGAGAGAGATTTCTGTAATCTTCCAGATTTTGGAAATCAGATGATG 1299
 OY 841 TTGGCTTGCATCATGTGAAAAGGATCAGAAATATTAAGTCTGTTGTAATTTAGAGAAA 900
 DB 1300 TTGGCTTGCATCATGTGAAAAGGATCAGAAATATTAAGTCTGTTGTAATTTAGAGAAA 1359
 OY 901 TCGTTTGGAAATTA 915
 DB 1360 TCGTTTGGAAATTA 1374

RESULT 2
 AA92646
 ID AA92646 standard; cDNA; 2822 BP.

XX AA92646;
 XX 05-JUN-2000 (first entry)

XX Human alpha-1,3-fucosyltransferase cDNA, SEQ ID NO.4.
 XX
 XX Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope;
 KW brain; kidney; recombinant expression; transgenic animal; knockout
 KW animal; FUC-TV; drug screening; inhibitor; potentiators; diagnosis;
 KW treatment; cancer; human; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 XX FT 1289..2368
 XX CD5 /tag-
 FT /product- "Human alpha-1,3-fucosyltransferase"

PN MO200006708-A1.

PD 10-FEB-2000.

PF 29-JUL-1999; 99WO-JP04092.

PR 29-JUL-1998; 98JP-0213823.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Narimatsu H, Kudo T, Sasaki K;

DR WPI: 2000-183120/16.
 DR P-PSDB: AAY80996.

invention also relates to DNA sequences encoding alpha-1,3-fucosyltransferase and expression vectors and host cells comprising these DNA sequences. The invention additionally encompasses the preparation of alpha-1,3-fucosyltransferase via the culture of transformed cells or by expression of the protein in a transgenic animal; antibodies which recognize alpha-1,3-fucosyltransferase; methods for screening potential inhibitors or potentiators of alpha-1,3-fucosyltransferase activity or expression; the preparation of compounds having fucose-containing sugar chains by use of the protein; and knockout non-human animals lacking alpha-1,3-fucosyltransferase. Alpha-1,3-fucosyltransferase has a similar substrate range to the known FUC-TIV and is expressed mainly in brain and kidney tissues. Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies, potentiators and inhibitors may be used for the treatment and diagnosis of diseases of the brain and kidney, and of cancers. They may be used for the identification of substances which affect the activity or expression of alpha-1,3-fucosyltransferase; such substances may be used therapeutically. The knockout animals can be used to study the mechanisms of action and expression of alpha-1,3-fucosyltransferase. Sequence AA292645 represents cDNA encoding murine alpha-1,3-fucosyltransferase (AA292645), and sequences AA292670-292673 are murine alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid amplification of cDNA ends).

Sequence 2036 BP; 666 A; 365 C; 348 G; 657 T; 0 other;

Query Match 86.5%; Score 791.8; DB 21; Length 2036;

Best Local Similarity 91.6%; Pred. No. 3.6e-219; Matches 838; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

OY 1 ACCAAAGTATTTTAAATGAACACTATCTGTGTGGTGGTGGCCATTGGGCGAG 60
DB 174 AAAAAAGATATTTTAAAGAAATACCATCTGTGGTGGTGAATGGCCATTGGGCGAG 233
OY 61 ACCCTTGACCTTACATCTCTGCGAACCAATGTTCACATCCAAAGATGCCATCTGACAACG 120
DB 234 ACCCTTGACCTTACATCTCTGCGAACCAATGTTCACATCCAAAGATGCCATCTGACAACG 293
OY 121 GACCGTCACGTGTACAAACAAATCCCATGAGTTCATCCATCCAGGACATCATGAGTGG 180
DB 294 GACCGTCACGTGTGTACAAACAAATCCCATGAGTTCATCCATCCAGGACATCATGAGTGG 353
OY 181 GATCGACCAATTTTACCTGAGCAAGTACGACCTTCCGAAATGGATTTGATGATGAT 240
DB 354 GATCGACCAATTTTACCTGAGCAAGTACGACCTTCCGAAATGGATTTGATGATGAT 413
OY 241 TTGGATGACCAATCTACACTCCCAAAAGATGGCATTTGATGATGATGAT 300
DB 414 TTGGATGACCAATCTACACTCCCAAAAGATGGCATTTGATGATGATGATGAT 473
OY 301 CTGACTTACCGCGGTGATTCAGATATTCAGATGCTTATGCGCTTGTAGCGTAAACCA 360
DB 474 CTGACTTACCGCGGTGATTCAGATATTCAGATGCTTATGCGCTTGTAGCGTAAACCA 533
OY 361 AATCCCTTCGTGTGAGTGGCAAGCAAGAAATTTGATGCTGGTGTGATGATAC 420
DB 534 AATCCCTTCGTGTGAGTGGCAAGCAAGAAATTTGATGCTGGTGTGATGATAC 593
OY 421 TGGAAACCTGAGCATGCGAGATCAAGTATTACATGATGACCAAGCAAGATGGAATC 480
DB 594 TGGAAACCTGAGCATGCGAGATCAAGTATTACATGATGACCAAGCAAGATGGAATC 653
OY 481 CATPACTACGGGCAAGCAATTTGGAGATATGATGATGATGATGATGATGATGAT 540
DB 654 CATPACTACGGGCAAGCAATTTGGAGATATGATGATGATGATGATGATGATGAT 713
OY 541 TCTGCTGTAATTTTATCTTCTTGAATTAATCAATCCCAAGATATCATCAGGAA 600
DB 714 TCTGCTGTAATTTTATCTTCTTGAATTAATCAATCCCAAGATATCATCAGGAA 773
OY 601 AAGCTATACATGCTTTCTGCGTGGCTGCTGATCCTGTTGTTCTGGACCATCTAGGAA 660
DB 774 AAGCTATACATGCTTTCTGCGTGGCTGCTGATCCTGTTGTTCTGGACCATCTAGGAA 833

```

```

OY 661 AACTATGAGATTTATATTCACAGATTCATTCATTCATGATGAGATTTAACTCTCC 720
DB 834 AACTATGAGATTTATATTCACAGATTCATTCATTCATGATGAGATTTAACTCTCC 893
OY 721 AGTAGAGTACCAAGATATTCAGAGAGATGACAAAAACATATAGTTATACCTTAGTAC 780
DB 894 AGTAGAGTACCAAGATATTCAGAGAGATGACAAAAACATATAGTTATACCTTAGTAC 953
OY 781 TTTAACTGAGAGAGATTTCACTGTAATCTTCCACGATTTTGGGAATCACATGATCT 840
DB 954 TTTAACTGAGAGAGATTTCACTGTAATCTTCCACGATTTTGGGAATCACATGATCT 1013
OY 841 TTGCTTCGATCATGATGAAAGAGCATCAAGATATATAGTCTGTTATTTAGAGAA 900
DB 1014 CTGGCATGCGACCATGTAATAAAGCATCAAGATATATAGTCTGTTATTTAGAGAA 1073
OY 901 TGGTTTGGGAATTA 915
DB 1074 TGGTTTGGGAATTA 1088

```

RESULT 4

AA292645
ID AA292645 standard; cDNA; 2170 BP.

AA292645;

05-JUN-2000 (first entry)

Murine alpha-1,3-fucosyltransferase cDNA.

Alpha-1,3-fucosyltransferase; fucose; glycosylation; Lewis epitope; brain; kidney; recombinant expression; transgenic animal; knockout animal; FUC-TIV; drug screening; inhibitor; potentiators; diagnosis; treatment; cancer; murine; mouse; ss.

Mus sp.

Key Location/Qualifiers

CDS 115..1194 /tag="a /product="Murine alpha-1,3-fucosyltransferase"

W0200006708-A1.

10-FEB-2000.

29-JUL-1999; 99MO-JF04092.

29-JUL-1998; 98JP-0213823.

(RYOW) KYOWA HAKKO KOGYO KK.

Narimatsu H, Kudo T, Sasaki K;

WPI; 2000-183120/16.

P-SDB; AA290995.

Alpha 1,3-fucosyltransferase generating Lewis x but not sialyl-Lewis x epitope and an antibody recognizing it useful for diagnosis of brain and kidney disease and cancer.

Claim 4; Page 127-134; 172pp; Japanese.

The invention relates to a novel alpha-1,3-fucosyltransferase which transfers a fucose moiety to galactosyl-beta-1,4-N-acetylglucosamine (generating the Lewis x or y epitope). It does not transfer a fucose moiety to alpha-2,3-sialyl-galactosyl-beta-1,4-N-acetylglucosamine and therefore does not generate the sialyl-Lewis x epitope. The invention also relates to DNA sequences encoding alpha-1,3-fucosyltransferase and expression vectors and host cells comprising these DNA sequences. The invention additionally encompasses the

CC preparation of alpha-1,3-fucosyltransferase via the culture of
 CC transformed cells or by expression of the protein in a transgenic animal;
 CC antibodies which recognise alpha-1,3-fucosyltransferase; methods for
 CC screening potential inhibitors or potentiators of
 CC alpha-1,3-fucosyltransferase activity or expression; the preparation of
 CC compounds having fucose-containing sugar chains by use of the protein;
 CC and knockout non-human animals lacking alpha-1,3-fucosyltransferase.
 CC Alpha-1,3-fucosyltransferase has a similar substrate range to the known
 CC Fuc-TIV and is expressed mainly in brain and kidney tissues.
 CC Alpha-1,3-fucosyltransferase, nucleotides which encode it, antibodies,
 CC potentiators and inhibitors may be used for the treatment and diagnosis
 CC of diseases of the brain and kidney, and of cancers. They may be used for
 CC the identification of substances which affect the activity or expression
 CC of alpha-1,3-fucosyltransferase; such substances may be used
 CC therapeutically. The knockout animals can be used to study the mechanisms
 CC of action and expression of alpha-1,3-fucosyltransferase. Sequence
 CC AA292645 represents cDNA encoding murine alpha-1,3-fucosyltransferase
 CC (AA80995), and sequences AA92670-292673 are murine
 CC alpha-1,3-fucosyltransferase exon 1 sequences obtained by 5' RACE (rapid
 CC amplification of cDNA ends).

XX Sequence 2170 BP; 715 A; 395 C; 368 G; 692 T; 0 other;

Query Match 86.5%; Score 791.8; DB 21; Length 2170;

Best Local Similarity 91.6%; Pred. No. 3.7e-219;

Matches 838; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 ACCAAACGATGATTTTAAATGAACACTGTTTGTGGTGGTGGCCATTGGGCGAG 60
 DB 280 ACAAAAGTGATTTTAAAGCAACATCCATCTGTGGTGGTGGCCATTGGGCGAG 339
 QY 61 ACCTTGAACCTTACATCCCTGGCAAGCAATGTTCAATCCAGAGATCCATTCACAACG 120
 DB 340 ACCTTGAACCTTACATCCCTGGCAAGCAATGTTCAATCCAGAGATCCATTCACAACG 399
 QY 121 GACGCTGACGTACCAACAAATCCAGTCTGATCCATCCAGACATCAGTTGG 180
 DB 400 GACGCTGATTTGACCAACAAATCCAGTCTGATCCATCCAGACATCAGTTGG 459
 QY 181 GATCTGACAAATTTTACCTGAGCAAGTACGACCTTCCAGAAATGGATTGGATGAT 240
 DB 460 GATCTGACAAATTTTACCTGAGCAAGTACGACCTTCCAGAAATGGATTGGATGAT 519
 QY 241 TTGGAATCACAACCTCAACCTCCCAAAAGAGTGGCATTTGACACTTGAACCTGACT 300
 DB 520 TTGAGAGTCAACCTCAACCTCCCAAAAGAGTGGCATTTGACACTTGAACCTGACT 579
 QY 301 CTGACTTACCGCGCGGTGATGATATCCAGTGGCTTATGGCTTCTGAGCGTAAAGACA 360
 DB 580 CTGACTTATCGCGCGGTGATGATATCCAGTGGCTTATGGCTTCTGAGCGTAAAGACA 639
 QY 361 AATCCCTTGGTGGTGAAGTCCCAAGCAAGAAATGGTGGTGGTGGTGAATGAT 420
 DB 640 AATCCCTTGGTGGTGAAGTCCCAAGCAAGAAATGGTGGTGGTGGTGAATGAT 699
 QY 421 TGGAAACCTGAGATGCGCAGATCAAGTATTACAAATGAGTAAAGCAAAAGCATTTGAATTC 480
 DB 700 TGGAAACCTGAGATGCGCAGATCAAGTATTACAAATGAGTAAAGCAAAAGCATTTGAATTC 759
 QY 481 CATACCTACGGGCAAGATTTGGAGATATGTCATGATATAAATTTGATTCCTACATA 540
 DB 760 CATACCTACGGGCAAGATTTGGAGATATGTCATGATATAAATTTGATTCCTACATA 819
 QY 541 TCTGCTTGAATTTTATCTTCTTGAATTCATTCACAGAGATTTACATCAGGAA 600
 DB 820 TCTGCTTGAATTTTATCTTCTTGAATTCATTCACAGAGATTTACATCAGGAA 879
 QY 601 AAGCTATACAAATGCTTTTGGCTGCTGCTGATCCTGTTGTTGGACCATCTAGGAGAA 660
 DB 880 AAGCTATACAAATGCTTTTGGCTGCTGCTGATCCTGTTGTTGGACCATCTAGGAGAA 939
 QY 661 AAGCTATACAAATGCTTTTGGCTGCTGCTGATCCTGTTGTTGGACCATCTAGGAGAA 939
 DB 661 AAGCTATACAAATGCTTTTGGCTGCTGCTGATCCTGTTGTTGGACCATCTAGGAGAA 939

DB 940 AACATGAGATTTATTTCCAGCTGATTCATTCATTTGAGCAAGATTTTAACTCTCCC 999
 QY 721 AGTAGAGTACCAAGATTTCTGAAAGAGTGCACAAAACATAAAGTATACCTTAGTAC 780
 DB 1000 AGTAGAGTACCAAGATTTCTGAAAGAGTGCACAAAACATAAAGTATACCTTAGTAC 1059
 QY 781 TTTAATCTGAGGAAGATTTTACCTGTTAAATCTTCCAGATTTTGGGAATCAGATGAT 840
 DB 1060 TTTAATCTGAGGAAGATTTTACCTGTTAAATCTTCCAGATTTTGGGAATCAGATGAT 1119
 QY 841 TTGCTTGGCATCTAGTGAAGAGCATCAAGATATATAGTCTGTGTAATTTAGAGAAA 900
 DB 1120 TTGCTTGGCATCTAGTGAAGAGCATCAAGATATATAGTCTGTGTAATTTAGAGAAA 1179
 QY 901 TGGTTTGGAAATTA 915
 DB 1180 TGGTTTGGAAATTA 1194

RESULT 5

AAT59505 standard; cDNA; 1814 BP.

AC AAT59505;

DT 06-MAY-1997 (first entry)

DE Murine myeloid-1 lineage alpha-(1,3)-fucosyltransferase cDNA.

KW Alpha-(1,3)-fucosyltransferase; fucosylation; antibody; IgG; IgM;

XX septic shock; septicemia; therapy; ss.

OS Mus sp.

FT Key location/Qualifiers

FT CDS 325..1353

FT /tag= a

PN WO9640881-A1.

PD 19-DEC-1996.

PF 08-MAY-1996; 96WO-0506427.

PR 07-JUN-1995; 95US-0483151.

PA (GENE) GEN HOSPITAL CORP.

PI Holgersson J, Seed B;

DR WPI: 1997-108639/10.

DR P-PSDB; AAM11820.

PT New murine alpha-(1,3)-fucosyltransferase - for fucosylating an

PT antibody to protect mammals against e.g. septic shock or septicemia

PS Claim 6; Fig 3; 58pp; English.

XX A cDNA clone (AAT59505) codes for murine myeloid lineage alpha-

CC (1,3)-fucosyltransferase (AAM11820), an enzyme which has a relatively

CC strict substrate requirement for sialylated N-acetylglucosamine,

CC which can account for the presence of the sialyl-Le^x epitope on

CC murine cells, and which is more effective than Fuc-TIV in support

CC of E-selectin-mediated COS cell adhesion. It was isolated from a

CC murine myeloid 32D cl3 cell cDNA library by identification of a

CC clone capable of directing the expression of sialyl-Le^x

CC determinants. Transformed host cells (e.g. 32D cl3 or human 293

CC cells) expressing the alpha-(1,3)-fucosyltransferase, and pref.

CC another fucosyltransferase such as human Fuc-TIV (see also AAT59506),

CC can be used to fucosylate an antibody (e.g. IgG or IgM) for use in

CC protecting an animal against an adverse immune reaction, esp. septic

CC shock or septicemia.

Tue Oct 8 10:18:37 2002

us-09-744-748-5_copy_460_1374.rng

Page 6

SO Sequence 1814 BP; 360 A; 523 C; 518 G; 413 T; 0 other;

Query Match 13.9%; Score 127.4; DB 18; Length 1814;

Best Local Similarity 52.18; Pred. No. 1.3e-26; Mismatches 334; Conservative 0; Indels 6; Gaps 2;

104 GATGCCATCTACAAAGGACCGTTCCTGTAACAACAAATCCCATGCTGATCCATC 163
104 GATGCCATCTACAAAGGACCGTTCCTGTAACAACAAATCCCATGCTGATCCATC 163
548 GCTGCGCTGCTAGTCTACCGGAGCGCTGAGTCAAGTCTGCTGCTCCACG 607
164 ACCGAGCATGAGTTGGGATCTGACAAATTTACCTCAGCAAGCTAGGCGCTTCGAGA 223
608 ACCGAGCATGAGTTGGGATCTGACAAATTTACCTCAGCAAGCTAGGCGCTTCGAGA 223
224 AATGATTTGGATGAAATTTGGAAATCCCACTCAGCAAGCTAGGCGCTTCGAGA 283
668 CTTGGCTGCTGGGCTCCATGGAATGCGGCAATGAGTCAATGCTGCTGCTGCTG 727
284 ACTGCTTAACTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 343
728 GATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 787
344 TCTGAGCGTAAAGCAAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
788 GCTTGAGACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 844
404 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
845 CTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 904
464 GCAAAAGCATGAAATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
905 CTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
524 ATTGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
965 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
584 AGATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 640
1025 GGGATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1084
641 TTTGAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 700
1085 CGCTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1144
701 TGGAGATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 741
1145 TGGAGATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1185

RESULT 6
AAT85219
ID AAT85219 standard; DNA; 3594 BP.

AC AAT85219;

XX 02-MAR-1998 (first entry)

XX Mouse alpha-fucosyltransferase Fuc-TVIII gene.

XX Alpha-fucosyltransferase; Fuc-TVII gene; mouse;

KW slaly1 Lewis x determinant; oligosaccharide; antiinflammatory;

KW inhibitor; de.

XX Mus musculus.

XX Mus musculus.

XX Key 996.3082

FT exon /note- "includes introns"

FT /tag- b

FT /number- 1
FT 478..669
FT /tag- c
FT /number- 1
FT 670..1149
FT /tag- d
FT /number- 2
FT /codon_start= 996..998
FT 1150..1946
FT /tag- e
FT /number- 2
FT 1947..1959
FT /tag- f
FT /number- 3
FT /codon_start= 1947..1949
FT /note- "alternative ATG start codon"
FT 1960..2065
FT /tag- g
FT /number- 3
FT 2067..3552
FT /tag- h
FT /number- 4
FT /codon_start= 2126..2128
FT /note- "alternative ATG start codon"

W09732889-A1.

PD 12-SEP-1997.

XX 07-MAR-1997; 97W0-DS03184.

XX 08-MAR-1996; 96US-0613098.

XX (UNMI) UNIV MICHIGAN.

XX Gersten KM, Lowe JB, Natsuka S;

XX WPI; 1997-470535/43.

XX P-PSDB; AAM27138 and AAM26671.

XX Novel murine alpha-(1-3)-fucosyl-transferase(s) - useful for e.g. screening for inhibitors used as antiinflammatory agents

XX Claim 1; Fig 2; 91pp; English.

CC This DNA sequence comprises the mouse Fuc-TVIII gene encoding
CC alpha-fucosyltransferase. It was isolated from a 373 cell genomic
CC library by screening with a probe from the catalytic domain of the
CC human Lewis alpha(1,3/4)fucosyltransferase (Fuc-TVII).
CC Examination of the mouse Fuc-TVIII locus identified 3 Met codons
CC that may initiate translation of alpha(1,3) fucosyltransferases
CC with different cytosolic domains encoded by exons 2 and/or 3, but
CC (see AAM27138 and AAM26671). The proteins and nucleic acids can
CC be used to construct animal cell lines able to post-translationally
CC modify oligosaccharides on cell surface, intracellular or secreted
CC proteins and lipids, to isolate reagents for efficient enzymatic
CC production of oligosaccharides, to generate antibodies to
CC glycosyltransferases useful as diagnostic reagents, to screen for
CC fucosyltransferase inhibitors and inactivators, especially those
CC that act as antiinflammatory agents, for genotyping individuals at
CC the fucosyltransferase locus and for in vitro synthesis of slaly1
CC Lewis x tetrasaccharide.

SO Sequence 3594 BP; 731 A; 969 C; 1067 G; 827 T; 0 other;

Query Match 13.9%; Score 127; DB 18; Length 3594;

Best Local Similarity 52.3%; Pred. No. 2.4e-26; Mismatches 295; Indels 6; Gaps 2;

104 GATGCCATCTACAAAGGACCGTTCCTGTAACAACAAATCCCATGCTGATCCATC 163
2277 GCTGCGCTGCTAGTCTAACCGGAGCGCTGTACCGAGTCTGATGCTGCTTCCACG 2336

```

QY 164 ACCGAGACATCAGTTGGGATCTGACAAATTTACCTCAGCAAGCTAGGCCACCCCTCCAGA 223
    |||||
DB 2337 ACCGTAGCTGGAACCCCGGCAATCTCTCCCTACCCCTGAGACAGGCGACAGGACAG 2396
QY 224 AATGATTTGGAGATTTTGAATCCCAACACACACACACACACACACACACACACACAC 283
    |||||
DB 2397 CTGGGGCTGGGCTCCACAGATCCCAAGTAATACCCATGGTCTCCATCGCTCCGGG 2456
QY 284 ACTGTTTAACTGACTGACTTACCGCGGTGATCAGATATCCAAATGAGCTTATGGCT 343
    |||||
DB 2457 GCATCTTCAACGTGGGCTGAGCTATCGCGGTGATTCATATCTTTTACCTTACGGTTC 2516
QY 344 TCTTGACGGTAAAGCACAATTCCTTCGTTGTTGAACTGGCCAAAGAGAAATTTGGTGT 403
    |||||
DB 2517 GCTTGAGGCTCTCTCTGGGCCC---CACATCCCACTACCGGCGCAAAAGACAGATGGCTG 2573
QY 404 GCTGGGTTTGAGTATCTGGAACCCCTGAGCATGCCAGACTCAAGTATTACATGAGCTAA 463
    |||||
DB 2574 CCGTGGGTATCAGCAATTTTCCAGGAGCGGCGACAGCGTCCAAAGCTACCGGCGACCTGG 2633
QY 464 GCAAAAAGCATTTAAATTCATACCTACGCGCAAGCATTTGGAGAAATATGCAATGATATAA 523
    |||||
DB 2634 CCCCTCATCTGAGGTGAGTGTGTTGCGCGCCAGCGGAGCGCCCTATGCGCTAATTT 2693
QY 524 ATTTGATTCCTACCATATCTGCTTTAAATTTTATCTTCTTGAATTAATTCATTCACACA 583
    |||||
DB 2694 GCTGCTGCGCCACTTTGGCCGCTGACGCTCTACCTGCGCTTTGAGAACTGACAGACATC 2753
QY 584 AGGATTTACATCAGGAGAAAGCTAT---ACAATGCTTTTCTGGCTGGCTCTGACCTGTTG 640
    |||||
DB 2754 GGGACTACATCAGTGAAGTCTGGCGCAATGCCCTGGCGGCTGGTGTGCTGACCCGCTGG 2813
QY 641 TTCTGGGACCATCTAGGAGAAATATGAGATTAATTTCCAGCAGATTCATTCATTCATG 700
    |||||
DB 2814 CCGTGGGACCTCTCTCGGCGCACCCTACGAGGCTTTTGTGCGCAGCATGCTTTGTACAGG 2873
QY 701 TGGAGATTATTAATCTCTCCAGTGAAGCTAGC 731
    |||||
DB 2874 TGGAGCATTTCAAGCTCTGCGCGTGAACCTGGC 2904

RESULT 7
AA077732
ID AA077732 standard; cDNA to mRNA; 1701 BP.
AC AA077732;
XX
XX 22-JUN-1995 (first entry)
DE Human alpha-1,3-fucosyltransferase cDNA.
XX
XX Alpha-1,3-fucosyltransferase; sialylated Lewis blood group antigen;
KM Lewis x; Lewis a; glycosylation; glycoprotein; glycolipid; ds.
XX
XX Homo sapiens (THP-1 cells).
OS
XX
XX Key Location/Qualifiers
FH 123..1151
FT CDS
FT /tag= a
FT /product= alpha-1,3-fucosyltransferase
XX
XX WO9423021-A.
PN 13-OCT-1994.
XX
XX 28-MAR-1994; 94WO-JP00496.
XX
XX 29-MAR-1993; 93JP-0069016.
XX
XX (KYOM ) KYOMA HAKKO KOGYO KK.
XX
XX Hanai N, Kurata K, Nishi T, Sasaki K;
PI

```

```

XX WPI: 1994-333183/41.
DR P-PSDB; AAR63215.
XX
XX Alpha-1,3-fucosyl-transferase DNA and protein - useful for prodn.
PT of active sugar chains, e.g. sialylated Lewis x
PS Claim 3; Page 99-102; 155pp; Japanese.
XX
XX The sequence coding for alpha-1,3-fucosyltransferase was isolated
CC from a human THP-1 cell cDNA library. The enzyme catalyses the
CC addition of fucose via a 1,3-linkage to N-acetylglucosamine
CC contained in glycoproteins and glycolipids, e.g. for sialylation of
CC Lewis blood group antigens.
XX
XX Sequence 1701 BP; 292 A; 524 C; 586 G; 299 T; 0 other;
SO
Query Match 12.5%; Score 114.4; DB 15; Length 1701;
Best Local Similarity 50.8%; Pred. No. 7,6e-23;
Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

QY 91 TTCACATCCAGGATGGCATCTCACAAGGACCCCTTCACTAGTACAAATCCCATGCA 150
    |||||
DB 333 TAGGCGATGCCCCGCTGCGCACCTGAGTGCACCAAGCCTGCTGGCCAGCGCGACGCC 392
QY 151 GTTCTGATTCACATCCCGAGACATCAATTTGATGACAAATTTACCTAGCAAGCTAGG 210
    |||||
DB 393 GGTGTCTTCCACACCGGAGAGCTGCAGACCCGGGSGTCCACCTGCGCCCGCCAGCGG 452
QY 211 CCACCTTCACAGAAATGATTTGGATGAATTTGAAATCAACAATCACAACCTCCCAAAAG 270
    |||||
DB 453 CCGCGAGGCGAGCCCGGGGTGGGCTCCATGAGTCTCTAGCCACACACCGACGCCCTC 512
QY 271 AGGCGATTTGAGACCTGTTTAACTGATCGATCGACTTACCGCGGATTCATATTCACA 330
    |||||
DB 513 AGCCACCTCCAGAGCAATCTTCAACTGGGTGCTAGCTTACCGCGGCACTGAGACATCTTT 572
QY 331 GTGCTTATGCTTCTTGAAGGTAAGCAAAATCCCTTGTGTTTGAATGCAAGCAAA 390
    |||||
DB 573 GTGCTTATGCGCGCTGAGAGCCCGCTGAGGCGCT---CGCACCGCTGCGACGCAAG 629
QY 391 GAGAAATGGTGTGCTGGGTTGTGAGTAACGAAACCTGAGCAATGCCAGAGTCAAGTAT 450
    |||||
DB 630 AGCAGGAGTGGCGCGCTGGGTGTGACCAACTTCCAGAGCGGAGGTGCTGCGAGGCTG 689
QY 451 TACATGAGCTTAAGCAAAAGCATTTGAATCCATACCTACGGCGCAAGCATTTGAGATAT 510
    |||||
DB 690 TACCGGACAGCTGGCGCTCATCTGCGGTGAGATGTCTTTGGCGCTGCCAATGAGCGGCA 749
QY 511 GTCAATGATTAATAATTTGATTCCTACCATATCTGCTTGAATTTTAAATTTTCTTGA 570
    |||||
DB 750 CTGTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
QY 571 AATTCATCCAGAGGATTTACATACAGGAAAGCTAT---ACAATGCTTTTCTGGCTGGC 627
    |||||
DB 810 AACTCTACAGCAGCGGACATACATTTAGGAGAAATTTGCGGCAAGCACTGGTGGCTGGC 869
QY 628 TCTGTACCTGTGTGTGTGGGACCATCTAGGAAATATGAGAAATATATTTCCAGAGAT 687
    |||||
DB 870 ACTGTGCGAGTGTGTGTGGGCGCCCGCCAGCGGCGCACTAGAGAGCCCTTCTGCGGCTGAC 929
QY 688 TCATTCATTCATGTGGAAGATTAATACTCCAGTGAAGCTAGC 731
    |||||
DB 930 GCCTTGTGATGTGATGATCTTTGGCTGACGCCGAGAGCTGGC 973

RESULT 8
AAF21136
ID AAF21136 standard; DNA; 1701 BP.
XX
XX AAF21136;
AC
XX 14-MAR-2001 (first entry)
DT

```

XX DE Human low adenosine antisense oligonucleotide related sequence #2703.
 XX XX human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary vasoconstriction; impaired respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX OS Homo sapiens.
 XX OS WO200062736-A2.
 XX PN 26-OCT-2000.
 XX PD 24-MAR-2000; 2000WO-US08020.
 XX PF 06-APR-1999; 99US-0127958.
 XX PR (UYEC-) UNIV EAST CAROLINA.
 XX PA (NYCE/) NYCE J W.
 XX PI Nyce JW;
 XX WPI: 2000-679539/66.
 XX DR Low adenosine (A) content antisense oligonucleotides which do not
 XX PT trigger adenosine receptors during metabolism, useful e.g. for treating
 XX PT cancers and respiratory obstructions -
 XX PS Disclosure: Page 1034; 1592pp; English.

CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and/or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptidic factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impaired respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 XX XX
 SQ Sequence 1701 BP: 290 A; 527 C; 589 G; 295 T; 0 other;

Query Match 12.5%; Score 114.4; DB 21; Length 1701;
 Best Local Similarity 50.8%; Pred. NO. 7.6e-23;
 Matches 327; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

OY 91 TTCACATCCAAAGATGCGATCTCACACGAGCGGTCACGTACACAAATCCCATGCA 150
 DB 333 TACGGATGCGCCCGCCGACCTGAGTGCCACCAAGCCCTGCGACGCGGACGCC 392
 OY 151 GTTGTGATCCATCACCAGACATCAGTTGGATGAGCAAAATTTACCTCAGACAGCTAGG 210
 DB 393 GTGGCTTTCACACCGGACGAGCTGACAGACCCGCGCTCCACCTGCCCCAGCGG 452
 OY 211 CCACCCCTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 270
 DB 453 CCGCGAGGCGACCCCTGGGTGTGGCCCTCCATGAGTCTCTAGCCACACCCAGCGGCTC 512
 OY 271 AGTGCATGACACACTGTTGTTAACTGATGATGATGATGATGATGATGATGATGATGAT 330
 DB 513 AGCCACCTCCGAGGATCTTCAACTGGGTGTGAGCTACCGCGGAGCTCGGACATCTTT 572
 OY 331 GTGCTTATGCGCTGCTGTGACGGTAAAGCAAAATCCCTGCTGTGAAAGTGCACAA 390
 DB 573 GTGCCCTATGCGCCGCTGAGGCCCACTGCGGCGCT---CGCACCGCTGCAAGC 629
 OY 391 GAGAAATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 450
 DB 630 AGCAGGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
 OY 451 TACAAATGACCTAAGCAAAAGCATTTGAAATCCATACCTACGCGCAAGCATTTGAGAA 510
 DB 690 TACCGGACAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 749
 OY 511 GTCAATGATAAATTTGATTTCTTACCATATCGCTGTTAAATTTATCTTCTTTGAA 570
 DB 750 CTGTGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
 OY 571 AATTCATCCACAGATATATCATCAGGAAAGCAT---ACAATGCTTCTGCTGCTGCTG 627
 DB 810 AACTCTCAGCAGCGGACCTACATTAAGGAAATTTCTGGCGCAACGCACTGTGCTGCTG 869
 OY 628 TCTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 687
 DB 870 ACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
 OY 688 TCATTCATTCATGCTGAGATATTAACCTCTCCAGTACG 731
 DB 930 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 973

RESULT 9
 AAA35014
 ID AAA35014 standard; DNA; 1701 BP.
 XX AAA35014;
 AC 28-JUL-2000 (first entry)
 XX XX
 DE Human adenosine receptor related polynucleotide SRO ID NO:2703.
 XX XX
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukemia; lymphoma; carcinoma; metastasis; ss.
 XX OS Homo sapiens.
 XX OS WO200009525-A2.
 XX PN 24-FEB-2000.
 XX PD 03-AUG-1999; 99WO-US17712.
 XX PF 03-AUG-1998; 98US-0095212.

Db	5576	TACGGCATGCGCCGGCTGCCACCTGAGTGGCAACCGAAGCTGCTG66CCAGCGCGAGCC	5635
Qy	151	GTTTGATTCATACCGAGACATCACTTGGGATCTGACAAATTTACCTCAGCAACTAG	210
Db	5636	GTGGTCTTCACACCGCGAGCGTGCAGACCGCGCGGTCCACCTGCCCCCTGGCCAGCG	5695
Qy	211	CCACCTTCGAGAAATGAGTTGGATGAAATTGGAAATCAGAACTCCACTCCCAAA	270
Db	5686	CCGCGAGGAGACCCCTGGGTGTGGCCCTCCATGGATCTCTCAGCCACACCCAGGGCTC	5755
Qy	271	AGTGGCATTGAGACCTGTTTAACTGACCTGACTTACCGCGTGATTCAGATTCAA	330
Db	5756	AGCCACTCCGAGGACATCTCAACTGGGTGTGAGTACCGCGGAGACTCGGACATCTTT	5815
Qy	331	GTGCTTATGGCTTTCTTGACGCTAAGCACAAATCCCTTCGTGTTGAAGTCCACAGCAA	390
Db	5816	GTGGCCCTATGGCGCGCTGGAGCCCACTGGGGCCCT---CCGCAACGCTGCGACCC	5872
Qy	391	GAGAAATGGTGTGCTGCTGGATGTGACTGGAACCCGAGCATGCGAGATCAAGAT	450
Db	5873	AGCAGGGTGGCCGCTGGGTGTGTGACAACTTCAGAGAGCGGACCTGCTGCCAGGCTG	5932
Qy	451	TACAAATGAGTAAAGCAAAAGCATTGAATTCATACCTACGGGCAAGATTGGAGAAAT	510
Db	5933	TACGGGAGCTGGCGCCTCATCTGCGGGTGGATGATGCTTTGGCGCGCAATGAGAGGCCA	5992
Qy	511	GTCAATATATAAAAATTGATTCCTACCATATCTGCTGTGAATTTTATCTTCCTTGAA	570
Db	5993	CTGTGCGCCAGCTGCTCGGTGGTCCCGCGGCGCACTACCGCTTACCTGTGCTTTTAG	6052
Qy	571	AATTCATCCACAAGATTATCATCAGGAAAGCTAT---ACAATGCTTTCTGGCTGGC	627
Db	6053	AACCTCAGACCGCGACTCATTTACGAGAAATTTGGCGCAACGCACTGTTGGCTGGC	6112
Qy	628	TCGTGCTCTGTGTTGTGGGACCATCATGAGGAAACATATAGAAATATATTCACAGCAT	687
Db	6113	ACTGTGCAATGTGGTGGGGGCCCCCAGCGGGCCACTATGAGAGCCTTGTCGGCGGTAC	6172
Qy	688	TCATCTATTCATGTTGGAAGTTATTAATCTGCCAGTGAACCTAGC	731
Db	6173	GCCCTTGTGATGTGAATGACTTTGGCTAGGCCCGAGACTTGGC	6216

RESULT 12	ID	AA056905	standard; DNA; 1654 BP.
XX	AA056905;		
XX	AC		
XX	AA056905;		
DT	26-JUL-1994	(first entry)	
XX	PCDML1-alpha-(1-3)Fuc-Tv1	genomic DNA.	
DE			
XX	Glycosyltransferase; fucosyltransferase; GDP-Fuc; In vitro; cell.		
KM	surface; oligosaccharide; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	128..1208	
FT		/tag- a	
XX			
PN	MO9402616-A.		
XX			
PD	03-FEB-1994.		
XX			
PF	20-JUL-1993;	93MO-US06703.	
XX			
PR	20-JUL-1992;	92US-0914281.	
XX			
PA	(UNMI) UNIV MICHIGAN.		
PI			
	Lowe JB;		

XX MP1: 1994-048874/06.
DR P-PSDB; AAR45933.
XX
XX DNA fragment encoding a glycosyltransferase - can be used for in
FT vitro reactions to modify cell surface oligosaccharide(s) e.g.
FT blood gp. determinants, to protect against transplant rejection
XX
PS Claim 3; Fig 7; 249pp; English.
XX
XX The sequence is that encoding human GDP-Fuc: (beta-D-gal (1-4))-D-
CC GLcNAc alpha (1-3)-Fucosyltransferase. The enzyme produced by the
CC DNA may be non glycosylated. This prevents premature loss of enzyme
CC activity. It can also be used in in vitro reactions to modify cell
CC surface oligosaccharide moets. e.g. blood group determinants.
CC See also AAO65906-12.
XX
S0 Sequence 1654 BP; 298 A; 557 C; 459 G; 340 T; 0 other;

Query Match	Similarity	12.4%	Score 113.4	DB 15	Length 1654
Best Local	Similarity	51.2%	Prod. No. 1.5e-22		
Matches	Conservative	0	Mismatches 361	Indels	Gaps
Qy	11	ATTCTGGTGGTGGTGGCCATTCTGGGCGAGACCTTGGACCTTACATCCGACCGCAAGT	90		
Db	318	ATTCCTGCTGTGGAGCTGGCCCTTTTAACAACCCATTAAGCTCTGCCCGCTGCTCAGATG	377		
Qy	91	TT--CAACATCCAGATGCCATCTCACACGACCGCTTACTCTGTACAACAATCCAT	147		
Db	378	GTGCTGGACAGGCTGACTGCAACATCACTGCGACCGGAAGGTGTATCCAGAGGAGAC	437		
Qy	148	GCATCTTGATTCATACCCGAGACATCACTTGGATCTGA---CAATTTACTTAGCA	204		
Db	438	GGGTCATCTGTGCACACCGAGAGGTCTGTACAAACCCAGTGGCCAGCTCCACGCTCC	497		
Qy	205	GCTAGGCAACCTTCCAGAAATGGATTTGGATTTGGATTTGGATTCACAACTACACTCC	264		
Db	498	CCGAGGCGGAGGCGACAGATGATCTGTGTTCACATATGAAATCCCAAGCACTGCTGG	557		
Qy	265	CAAAAGATGTGACATGACACCTTGTAAACCGACTGCACTTACCGCGGTATTCAGAT	324		
Db	558	CAGCTGAAGGCATGAGAGGATCTTCAATCTCACATATCTCTACCGACGACCTCCGAC	617		
Qy	325	ATCCAATGCTTATGGCT-----TCTTACGCGTAAGCACAATCCCTGTGTTT	375		
Db	618	ATCTTCACAGCCCTACGAGCTGGCTGGAGCCGATGTCCGGACAGCTGCGCACCCAGCGCTC	677		
Qy	376	GAAAGTCCAAAGCAAGAAATTTGGTGTGTGTGGTGTGTGATTAAGTGAACCTTGACAT	435		
Db	678	AACCTCTGGGCGCAACGAGAGCTGTGGCTGTGGAGTGTCCAACTGGGGGCAAACTCC	737		
Qy	436	GCCAGATCAAGTATTTACAAATGAGCTAAGCAAAACATTTGAATTCATCTACCTACGAGCA	495		
Db	738	GCGAGGGTGGCTACTTACCAGAGCTGCGAGCCCATCTCMAAGTGGAGCTGTACGG--A	794		
Qy	496	GCATTTGAGATATATGATCAATGATTAATAATTTGATTTCTTACCATATCTGTTAAATTT	555		
Db	795	CGCTCCCAAGCCCTTGGCCGAGGAACCAATGATGAGAGAGCTGTCCCGGTACAAATTC	854		
Qy	556	TATCTTCTCTTGAATAATTCAATCCACAGAGATTACATCAGGAAAGCTAT--ACAAT	612		
Db	855	TATCTGGCTTTCGAAATCTCTTGGACCCCGGACTCATATCACCAGAAAGCTGTGAGGAAAC	914		
Qy	613	GCTTTTCGGCTGCTGTACTCTTGTCTTGGAGACATCTAGGGAACCACTATAGAAAT	672		
Db	915	GCCCTGGAGGCTGGGCCGTGGCCGTGTGTGTGTGGGCCCCAGCAGAAAGCACTACAGAGG	974		
Qy	673	TATATTCAGCAGATTCATTCATGTATGAGAAATTTAACTCTCCAGTAGTAGCA	732		
Db	975	TTCCGCGCACCGAGCGCTTATCTACAGTGGACGACTTCAGAGCCCAAGAGACTGGCC	1034		
Qy	733	AAGTATCTGAAGGAAGTGCACAAAACAATAAGTTATCTAGTACTTTAACTGAGG	792		

```

Db 1035 CGGTACCTGCAGAGCTGGACACAGAGCAGCCCGCTACTGAGCTACTTTCGTCGCG 1094
Qy 793 AAG 795
Db 1095 GAG 1097

RESULT 13
AAT61680
ID AAT61680 standard; DNA: 1654 BP.
XX
XX AAT61680;
AC
XX 19-JUN-1997 (first entry)
DE Human alpha(1,3)-fucosyltransferase (Fuc-TVI) DNA.
XX
XX Alpha(1,3/1,4)-fucosyltransferase; Lewis enzyme; Fuc-TVI;
KW glycosylation; oligosaccharide; blood group; ss.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 129..1208
FT /tag= a
PN MO9709421-A1.
PD 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-US13816.
PF
XX 08-SEP-1995; 95US-0525058.
PR
XX (UNMI ) UNIV MICHIGAN.
PA
XX Legault DJ, Lowe JB;
PI
XX MPI; 1997-192897/17.
DR P-PSDB; AAW13643.
XX
XX New recombinant fucosyltransferase proteins - useful for modifying
XX cell surface oligosaccharide structures
XX
XX Example 6; Page 291-292; 329pp; English.
XX
XX A DNA clone (AAT61680) codes for an enzyme (AAW13643) that functions as
XX a GDP-Fuc:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-fucosyltransferase
XX (Fuc-TVI) which can utilize neutral type II oligosaccharide
XX precursors but not type I glycoconjugates. It was identified in a
XX human lambda phage genomic DNA library using a probe obid. from
XX Lewis enzyme cDNA (see also AAT61675). The Fuc-TVI DNA can be used
XX to construct animal cell lines with specific capabilities with
XX respect to post-translational modification of the oligosaccharides
XX of expressed proteins or lipids, to produce recombinant Fuc-TVI for
XX use in oligosaccharide prodn., to genotype individuals at this
XX fucosyltransferase locus, and in the prodn. of chimeric
XX fucosyltransferases (see also AAW14512-32).
XX
XX Sequence 1654 BP: 298 A; 557 C; 459 G; 340 T; 0 other;
SQ
Query Match 12.4%; Score 113.4; DB 18; Length 1654;
Best Local Similarity 51.2%; Pred. No. 1.5e-22;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;

Db 31 ATTCTGCTGGTGGTGGCCATTGGGCGAGACCTTGGACTTACATCCTGGCAGCAGATG 90
Qy 318 ATCTGCTGGTGGTGGCCATTGGGCGAGACCTTGGACTTACATCCTGGCAGCAGATG 90
Db 318 ATCTGCTGGTGGTGGCCATTGGGCGAGACCTTGGACTTACATCCTGGCAGCAGATG 377
Qy 91 TT---CAACATCCAGATGCGATCTCCAAACGAGCGCTTCACTGTACAAACAAATCCCAT 147
Db 378 GTGCTGCGACGGCTGACTGCAACATCAGTCCGAGCCGCAAGGTATCCACAGGAGAC 437

Db 148 GCAGTTCTGATCATGACCGGAGACATGCTGGGATCTGA---CAAAATTACCTGAGCA 204
Qy 438 GCGGTATGCTGTGCACACCGAGAGGTATGTATACAAACCCGAGTCCATCCACGCTCC 497
Db 205 GGTAGGCGAACCCCTTCCAGAAATGGATTTGATTTGAAATCAACACTACACTCC 264
Qy 498 CCGAGCGCGGAGGCGAGGATGATCTGTTCACATGATGATCCGACAGCGCTGCG 557
Db 265 CAAAGAGTGGCATTTGAGCACTTGTAACTGATGATGATTTACCGCCGATTCAGAT 324
Qy 558 CAGCTGAAAGCCATGAGAGGATCTTCAATCTGACATGCTCTTACGAGGACTCGAC 617
Db 325 ATCCAGTGGCTTATGCGT-----TCTGAGGCTAAGCACAATTCCTTGGTCTT 375
Qy 618 ATCTTACGCGCTTACGGCTGGCTGAGGCGCTGTCCGCGACGCTCCGACCGCTC 677
Db 376 GAAATGCGCAAGCAAGAAATTTGATGCTGCTGCTGATGATGATGATGATGATGAT 435
Qy 678 AACCTCTCGGCCAAGACCGAGCTGTGGCTGGCGAGTGTCCAACTGGGGGCAACTCC 737
Db 436 GCCAGAGTCAATGATTTACAAATGAGTAAAGCAAAAGCATTTAAATCAATCAATCAAT 495
Qy 738 GCCAGGCTGCGCTACTACAGAGCTGCAAGGCCATCTCAAGTGGAGCTGACG---A 794
Db 496 GCATTTGGAGATATGATCATGATTAATTAATTTGATTCCTACCATATGCTGTAAATTT 555
Qy 795 CGCTCCCAAGCCCTGCTCCGAGGAGCCATGATGAGAGCGTGTCCGCTACAACTTC 854
Db 556 TATCTTCTCTTGAATAATTCACCAAGGATTTACATCAGCAAAAGCTAT---ACAAT 612
Qy 855 TATCTGCTCTTGAATAATTCCTTCAACCCGATCAATCAACGAGAACTGTGAGAGCA 914
Db 613 GCTTTTCTGCTGCTGTGATCTGATCTGTGTTCTGTGGACCATCTAGGAAACTATGAGAT 672
Qy 915 GCGCTGAGAGGCTGTGGCGCTGCCGCTGCTGTGGCGCCCAAGAACCACTGAGAGG 974
Db 673 TATATTCAGCAGATATTCATTCATTCATGAGATTAATTCATTCATTCATTCATTCAT 732
Qy 975 TTCTGCTGAGCAGGAGCTGCTGATCAGTGGAGCAGCTTCCAGAGCCGAGAGGACTGGCC 1034
Db 733 AAGTATCTGAAGAGAGTGCAGCAAAACAAATAGTTATACCTTACTTACTGAGAG 792
Qy 1035 CGGTACCTGCAGAGCTGGACAAAGACCGCCGCTACTGAGCTTCTGCTGCGCG 1094
Db 793 AAG 795
Db 1095 GAG 1097

RESULT 14
AAO56912
ID AAO56912 standard; DNA: 1086 BP.
XX
XX AAO56912;
DE 26-JUL-1994 (first entry)
XX
XX Lewis blood group fucosyltransferase DNA.
KW Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
XX surface; oligosaccharide; ss.
XX
XX MO9402616-A.
XX
XX 03-FEB-1994.
XX
XX 20-JUL-1993; 93WO-US06703.
XX
XX 20-JUL-1992; 92US-0914281.
XX
XX (UNMI ) UNIV MICHIGAN.
PI Lowe JB;

```

```

XX WPI: 1994-048874/06.
XX
XX DNA fragment encoding a glycosyltransferase - can be used for in
XX vitro reactions to modify cell surface oligosaccharides) e.g.
XX blood gp. determinants, to protect against transplant rejection
XX
XX Disclosure: Fig 6; 249pp; English.
XX
XX The sequence is that encoding a Lewis blood group fucosyl
XX transferase (Fuc-III). The enzyme produced by the DNA may be
XX non glycosylated. This prevents premature loss of enzyme activity.
XX It can also be used in in vitro reactions to modify cell surface
XX oligosaccharide mois. e.g. blood group determinants.
XX See also AA056905-11.
XX
SQ Sequence 1086 BP; 203 A; 387 C; 298 G; 198 T; 0 other;

Query Match      12.3%; Score 112.2; DB 15; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.7e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY 106 TGCATCTCACAACGAGCGTTCTACTGTACAAACAAATCCCATGCTGATCCATCAC 165
DB 271 TGCACATCTACTGCCAGCCGCAAGGTGTACCCACAGGCGACAGGTCATCGTGCACAC 330
OY 166 CGAGACATCAGTTGGA---TCTGACAAATTTACTCTACAGCAAGCTAGGCCACCTTCAG 222
DB 331 TGGGATATCATGTCCAAACCTTAAGTACAGCCTCCACCTTCCCGGCGCGAGGGGAG 390
OY 223 AATGATTTGGATGATTTGGATGATGACACACTCACCCTCCCAAAAAGAGTGCATTGAG 282
DB 391 CGCTGATCTGTCTCACTTGAGGACCCCTCAACTGCGACAGCTGGAAAGCTCGGAC 450
OY 283 CACTTGTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
DB 451 AGATATTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 510
OY 343 TCTTGTACGCTAAGCAACAAATCCCTTCGT-----GTTGAACTGCGCAACAAAGAG 393
DB 511 TGGCTGAGACCGGTGTGTCGCGCCAGCTGCCACCCAGCTCAACCTCGCGGCAGAC 570
OY 394 AATTGGTGTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
DB 571 GAGCTGTGCTGGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
OY 454 AATGACCTAAGCAAAAGATGAAATCCATACCTACGCGCAAGCATTTGGAAATATGTC 513
DB 631 CAGAGCTCAGAGCTCATCTCAAGGTGACGTGTACGAGCGCTCCACAAGCCCTGCC 690
OY 514 AATGATMAAAATTTGATCTCTACCATATCTGCTTGAATTTTATCTTCTTTGAAAT 573
DB 691 AAGGGAGCA---TGATGAGACGCTGTCCGGTACAACTTCTACCTGGCTTCGAGAAC 747
OY 574 TCAATCCACAAGATTAATCAACGGAAGCTAT---ACAATGCTTTTCTGGCTGCTCT 630
DB 748 TCTCTGACACCCGACACTACACCGGAGAGCTGTGAGGAGAGCCCTGAGGCTGGGCC 807
OY 631 GATACGTTGTTCTGTGGACATCTAGGGAAGAACTATGAGAAATTAATTTCCAGAGATTCA 690
DB 808 GTGCGCGTGTGTGGGCCCCAGACAGCAAGCAACTACGAAAGTTCCTCGACCCGCGGCC 867
OY 691 TTCTATCATGTGGAAGATTATTAATCTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 750
DB 868 TTCTATCATGTGGAAGATTATTAATCTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 927
OY 751 GACAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 795
DB 928 GACAAAGACAGCCCGCTACCTGAGACTACTTTCGCTGGCGGAG 972

```

RESULT 15
AAT80111

```

ID AAT80111 standard; cDNA; 2042 BP.
XX
XX AAT80111;
AC
XX
XX 01-DEC-1997 (first entry)
DT
XX
XX Fucosyltransferase III coding sequence.
DE
XX
XX P-selectin glycoprotein ligand-1; PSGL-1; human; mucin-like protein;
XX calcium-dependent carbohydrate binding protein; platelet; endothelium;
XX thrombin; leukocyte; sulphated glycosylated peptide; o-glycan; therapy;
XX inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
XX disseminated intravascular coagulation; rheumatoid arthritis; antibody;
XX adult respiratory distress syndrome; tumour metastasis;
XX fucosyltransferase III; acetylglucosaminyltransferase;
XX autoimmune disorder; inflammatory disorder; ss.
XX
OS Homo sapiens.
XX
XX W09706176-A2.
PN
XX
XX 20-FEB-1997.
PD
XX
XX 02-AUG-1996; 96MO-US12820.
PF
XX
XX 17-MAY-1996; 96US-0649802.
PR
XX 03-AUG-1995; 95US-0510920.
PR
XX 15-MAY-1996; 96US-0017794.
XX
XX (OKLA ) UNIV OKLAHOMA STATE.
XX
XX Cummings RD, Mcever RP, Moore KL;
XX
XX WPI: 1997-154206/14.
DR
XX
XX New O-glycan(s) derived from P-selectin glyco:protein ligand
XX (PSGL)-1 - used for inhibiting binding of PSGL-1 to selectins, e.g.
XX for treating inflammatory or autoimmune disorders or tumours
XX
XX Disclosure: Page 80-81; 96pp; English.
XX
XX This sequence represents the coding sequence for the human
XX fucosyltransferase III. The encoded protein, and the core 2 beta1-6-N-
XX acetylglucosaminyltransferase encoded by AAT80112 are needed for the
XX expression of the human P-selectin glycoprotein ligand-1 (PSGL-1) in
XX mammalian cells. P-selectin is a calcium-dependent carbohydrate binding
XX protein expressed on the surfaces of activated platelets and endothelium
XX in response to thrombin and other agonists. PSGL-1 (see AAW26174) is a
XX high affinity P-selectin ligand produced by leukocytes. Binding of
XX P-selectin to PSGL-1 is calcium ion dependent and is abolished by
XX treatment of the ligand with sialidase. PSGL is a homodimer, with a
XX highly extended extracellular domain, which is a feature of mucin-like
XX proteins. PSGL-1 is targeted by the sulphated glycosylated peptides of
XX the invention (see AAW26175-W26188). PSGL-1 is also targeted by the
XX O-glycans, and O-glycanated products of the invention. The products are
XX used for inhibiting the binding of P-selectin and other selectins to
XX PSGL-1. They can be used for the treatment of inflammation, injury
XX resulting from ischaemia and reperfusion, bacterial sepsis and
XX disseminated intravascular coagulation, adult respiratory distress
XX syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and
XX other autoimmune or inflammatory disorders. The products can also be used
XX in the production of antibodies and in diagnostic applications.
XX
SQ Sequence 2042 BP; 413 A; 646 C; 575 G; 408 T; 0 other;

Query Match      12.3%; Score 112.2; DB 18; Length 2042;
Best Local Similarity 51.6%; Pred. No. 3.6e-22;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

OY 106 TGCATCTCACAACGAGCGTTCTACTGTACAAACAAATCCCATGCTGATCCATCAC 165
DB 343 TGCACATCTACTGCCAGCCGCAAGGTGTACCCACAGGCGACAGGTCATCGTGCACAC 402

```

OY 166 CGAGACATCGATTGGA---TCTGACAAATTACTGACGCAAGCTAGGCCACCCCTTCAG 222
DB 403 TGGGATATCATGTCACACCTTAAGTCAGCGCTCCACCTTCCCGAGGGCCGAGGGGACG 462
OY 223 AATGATTTTGGATGATTTGGAATCACCACACTCCCAAAAGAGTGGCATTGAG 282
DB 463 CGCTGGATCTGTTCAACTGTGAGCCACCCCTTAAGTCCAGCACCCTGGAAGCCCTGAGC 522
OY 283 CACTTGTTCCTGACTGACTGACTTACCGCCGATTTCAGATTCAGAGTCCCTTATGGC 342
DB 523 AGATCTTCAATCTCAACATGCTCTACCGGAGGACTCCGACATCTTCAAGCCCTACGGC 582
OY 343 TTCTTACGGTAAAGCAAAATCCCTGCT-----GTTTGAAGTCCCAAGCAAGAG 393
DB 583 TGGCTGGAGCCCTGCTGCTCGGCCAGCCTGCCACCCACCGCTCAACCTCTCGGCCAAGACC 642
OY 394 AATTTGGTGTGCTGGGTTTGGAGTAAGTGAACCTGAGCATGCCAGAGTCAAGTATTAC 453
DB 643 GAGCTGGTGGCTGGGCGGTGTCACACTGGAAGCCGACTCAGCCAGGGTGGCTACTAC 702
OY 454 AATGAGCTAAGCAAAAGCATGGAATCCATACGCGCAAGCAATTGGAGAAATATGTC 513
DB 703 CAGAGCCTGAGGCTCATCTCAAGGTGAGCGTGTACGAGCGTCCCAAGCCCTGCCC 762
OY 514 AATGATAAAAATTGATTTCTTACCATATCTGCTTGTAAATTTTATCTTCTTTGAAAAAT 573
DB 763 AAGGGGACCA---TGATGAGAGCGCTGCCGCTACAGTTCTACCTGGCCTTCGAGAAC 819
OY 574 TCATTCACAAAGATTCATACACGAAAGCTAT---ACAATGCTTTCTGGCTGGCTCT 630
DB 820 TCCTTGCACCCCACTACATACCGAGAAAGCTGAGAGAACCCCTGAGAGGCGCTGGCC 879
OY 631 GTACTGTGTGTTCTGGGACCATCTAGGAAAACTATGAGAAATATATATCCAGCAGATTCA 690
DB 880 GTGCCGCTGTGCTGGGCCCAAGAGAACTACGAGAGGTTCCTGCCACCCGACGCC 939
OY 691 TTCTATTCATGTGGAAGATTATTAACCTCTCCAGAGAGCTAGCAAGTATCTGAAGGAAGTC 750
DB 940 TTCTATTCAGTGGAGACTTCCAGAGCCCAAGAGACTGGCCCGGTACTGAGAGAGCTG 999
OY 751 GACAAAACAAATAGTTATACCTTACTTAACTGGAGAG 795
DB 1000 GACAAAGACCAAGCCCGCTTACTGAGCTACTTTCGCTGGGGGAG 1044

Search completed: October 5, 2002, 23:16:20
Job time : 144.682 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:40:30 ; Search time 766.105 Seconds
(without alignments)
16120.141 Million cell updates/sec

Title: us-09-744-748-5_COPY_460_1374

Perfect score: 1 accaaactgattatttaa.....agaatggttttggaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	573.8	62.7	697	9	BB397507 BB397507
2	492.8	53.9	607	10	BB102760 60288338
3	422	46.1	716	9	AV725071 AV725071
4	324.8	35.5	649	9	AV338075 AV338075
5	307.4	33.6	579	10	BM021081 1e75103.x
6	286	31.3	1101	12	CNS05957 CNS05957
7	184.2	20.1	570	10	BM082484 f125e08.x
8	177	19.3	1100	12	CNS05792 CNS05792
9	170.2	18.6	1016	12	CNS0429P Tetradon
10	161.4	17.6	844	12	CNS0188V Tetradon
11	157.4	17.2	887	12	CNS0200S Tetradon
12	155.4	17.0	651	10	BJ036855 BJ036855
13	151.6	16.6	666	10	BJ050300 BJ050300
14	151.2	16.5	711	12	CNS034XF AL228012 Tetradon
15	151	16.5	501	12	BA1073 BS-1052-B2-
16	146.6	16.0	640	10	BJ073227 BJ073227
17	145.2	15.9	648	9	BB595164 BB595164

18	136.2	14.9	199	9	BB398103	BB398103
19	133	14.5	625	9	BB649667	BB649667
20	123.8	13.5	767	9	BB640911	BB640911
21	123.4	13.5	674	10	BE605503	BE605503
22	120.6	13.2	971	12	CNS0471V	CNS0471V
23	120.4	13.2	615	10	BI326811	BI326811
24	117.8	12.9	1023	12	CNS02ANG	CNS02ANG
25	102.8	11.2	487	10	BE863396	BE863396
26	100.2	11.0	702	12	BJ043578	BJ043578
27	93.8	10.3	863	12	CNS02ROF	CNS02ROF
28	90.4	9.9	720	10	BF180380	BF180380
29	90	9.8	551	10	BJ035701	BJ035701
30	87.4	9.6	601	10	BJ036679	BJ036679
31	87.4	9.6	605	10	BJ028804	BJ028804
32	87	9.5	1056	12	CNS01V17	CNS01V17
33	86.8	9.5	304	9	AU081478	AU081478
34	84.2	9.2	638	10	BJ035697	BJ035697
35	83.6	9.1	1057	12	CNS03ADC	CNS03ADC
36	83	9.1	566	10	BJ090565	BJ090565
37	80.2	8.8	628	9	BB651409	BB651409
38	79.4	8.7	776	12	CNS03RIO	CNS03RIO
39	78	8.5	568	10	BM021317	BM021317
40	78	8.5	838	10	BG432536	BG432536
41	76.4	8.3	655	10	BG538740	BG538740
42	75	8.2	878	12	CNS02VID	CNS02VID
43	71.6	7.8	590	10	BM082755	BM082755
44	70.4	7.7	582	9	AW233458	AW233458
45	70	7.7	505	10	BJ029303	BJ029303

ALIGNMENTS

RESULT 1
LOCUS BB397507 697 bp mRNA linear EST 24-OCT-2001
DEFINITION BB397507 RIKEN full-length enriched, ES cells Mus musculus cDNA
clone C330004C19 3' similar to AB015426 Mus musculus Fut9 mRNA for
alpha,3-fucosyltransferase IX, mRNA sequence.

ACCESSION BB397507
VERSION BB397507
KEYWORDS
SOURCE BB397507.2 GI:16411808
ORGANISM

REFERENCE 1 (bases 1 to 697)
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
house mouse.
Mus musculus.

REFERENCE 1 (bases 1 to 697)
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
house mouse.
Mus musculus.

REFERENCE 1 (bases 1 to 697)
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
house mouse.
Mus musculus.

REFERENCE 1 (bases 1 to 697)
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
house mouse.
Mus musculus.

REFERENCE 1 (bases 1 to 697)
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
house mouse.
Mus musculus.

QY 58 CAGACCTTTGACCTTACATCTCTGCCAAGCAATGTTCAACATCCAGAGATGCCATCTCACA 117
117
Db 1 CAGACCTTTGACCTTACATCTCTGCCAAGCAATGTTCAACATCCAGAGATGCCATCTCACA 60
QY 118 ACGAGACGGTTCACGTTCACAAACAAATCCCATGCAATTTGATCCATCAGCCAGACATCAGT 177
177
Db 61 ACAGACGGCTCA-TGTTCACAAACAAATCCCATGCGGTCTGTATCTCAGTAGACATCAGC 119
QY 178 TGGGATCTGACAAATTTACCTCAGCAGCAAGTAGGCCACCTCTCCAGAAATGGATTGGATG 237
237
Db 120 TGGGATCTGACAAATTTACCTCAGCAGCAAGTAGGCCACCTCTCTTCCAGAAATGGA-TTGGATG 178
QY 238 AATTGGAATCACCACCACTCAGCTCCCAAAAGAGTGGATGAGCACTTGTTCACCTG 297
297
Db 179 AATTGAGAGTACCCACCTCAGCTCCCAAAAGAGTGGATGAGCACTTGTTCACCTG 238
QY 298 ACCTGCTACACCGCCCTGATTCAGATATCCAGTGGCTTATGGCTTCTTGACGGTAAAGC 357
357
Db 239 ACCTGCTACACCGCCCTGATTCAGATATCCAGTGGCTTATGGCTTCTTGACGGTAAAGC 298
QY 358 ACAAAATCCCTGCTGCTTGAAGTCCCAAGCAAGAAATTTGGTGGCTGGTGTGAGT 417
417
Db 299 ACAAAATCCCTTGTGTTGAAGTCCCAAGCAAGAAATTTGGTGGCTGGTGTGAGT 358
QY 418 AACTGGAACCTGAGATGCGAGAGTCAAGTATTACAATGAGTAAACCAAAAGCATTGAA 477
477
Db 359 AACTGGAACCTGAGATGCGAGAGTCAAGTATTACAAGCAGTCAAGCAAGATGAA 418
QY 478 ATCATACCTACAGGGGCAAGATTTGGAGAAATATGTCAATGATTAATAATTTGATTCCTAC 537
537
Db 419 ATCCACACCTACAGGGGCAAGATTTGGAGAAATATGTCAAGTAAATAATTTGATTCCTAC 478
QY 538 ATATCTGCTGTAATTTTATCTTCTTGAATTCATCAATCCACAGAGATTTACATCAGC 597
597
Db 479 ATATCTGCTGTAATTTTATCTTCTTGAATTCATCAATCCACAGAGATTTACATCAGC 538
QY 598 GAAAAGCTATACATGCTTTCTGGGCTGCTGTGATCTGTTGTTGGACATCTAGG 657
657
Db 539 GAAAAGCTATACATGCTTTCTGGGCTGCTGTGATCTGTTGTTGGACATCTAGG 598
QY 658 GAAACACT 666
666
Db 599 GAAACACT 607
607

RESULT 3
AV725071 716 bp mRNA linear EST 16-OCT-2000
LOCUS AV725071 HTC Homo sapiens cDNA clone HTCLB02.5', mRNA sequence.
ACCESSION AV725071
VERSION AV725071.1 GI:10830091
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 716)
AUTHORS Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu
S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
Chen,J., Chen,Z. and Han,Z.
Homo sapiens cDNA HTC clones
COMMENT Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers

source 1. 716
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTCLB02"
/clone_1db="HTC"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 219 a 113 c 133 g 249 t 2 others
ORIGIN

Query Match 46.1%; Score 422; DB 9; Length 716;
Best Local Similarity 100.0%; Pred. No. 5e-106;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 AAGCATTTGGAGATATGTCATATATAAATTTGATTCCTACCATATCTGCTGTAAT 553
553
Db 1 AAGCATTTGGAGATATGTCATATATAAATTTGATTCCTACCATATCTGCTGTAAT 60
QY 554 TTTATCTTCTTCCGTAATTCATCCACAGATTTACATCAGGAAAGCTATACATG 613
613
Db 61 TTTATCTTCTTCCGTAATTCATCCACAGATTTACATCAGGAAAGCTATACATG 120
QY 614 CTTTCTGCGTGGCTCTGTAATTCATCTGTTGCGGACCATGAGGAAACATGAGATT 673
673
Db 121 CTTTCTGCGTGGCTCTGTAATTCATCTGTTGCGGACCATGAGGAAACATGAGATT 180
QY 674 ATATTCACAGCATTCATTCATTCATGAGAAATTAATCTCCAGTAGAGTACAA 733
733
Db 181 ATATTCACAGCATTCATTCATTCATGAGAAATTAATCTCCAGTAGAGTACAA 240
QY 734 AGTATCCGAGGAGAGTCCGCAAAACAAATAGTATACCTAGTACTTAACTGAGGA 793
793
Db 241 AGTATCCGAGGAGAGTCCGCAAAACAAATAGTATACCTAGTACTTAACTGAGGA 300
QY 794 AGGATTTCACTGTAAATCTTCCACAGATTTTGGAGATCAGATGTTGGCTTGCATC 853
853
Db 301 AGGATTTCACTGTAAATCTTCCACAGATTTTGGAGATCAGATGTTGGCTTGCATC 360
QY 854 ATGTGAAAGGCGATCAAGATATTAAGTCTGTGTAATTTAGAGAAATGGTTGGAAT 913
913
Db 361 ATGTGAAAGGCGATCAAGATATTAAGTCTGTGTAATTTAGAGAAATGGTTGGAAT 420
QY 914 AA 915
915
Db 421 AA 422
422

RESULT 4
AV338075 649 bp mRNA linear EST 24-OCT-2001
LOCUS AV338075
DEFINITION AV338075 RIKEN full-length enriched, adult male olfactory bulb Mus
musculus cDNA clone 6430407H24.3' similar to Ab015426 Mus musculus
mRNA for alpha1,3-fucosyltransferase IX, mRNA sequence.
ACCESSION AV338075
VERSION AV338075.2 GI:16396626
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 649)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muranatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
COMMENT Unpublished (2001)
TITLE
JOURNAL

COMMENT

On Nov 11, 1999 this sequence version replaced g1:6378127
Contact: Yoshihide Hayashizaki

Genetic Engineering Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-2 Suenori-cho
Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
E-mail: genpomo@riken.go.jp

REFERENCE	1 (bases 1 to 1100) Roest Croillius, H., Jalllon, O., Dasilva, C., Bouneau, L., Fisher, C., Benoit, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissendbach, J.
AUTHORS	
TITLE	Estimate of human gene number provided by genome-wide analysis using Tetradodon albigolyridis DNA sequence
JOURNAL	Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE	20296633
REFERENCE	2 (bases 1 to 1100) Croillius, H., R., Jalllon, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Benoit, A. and Weissendbach, J.
AUTHORS	
TITLE	Characterization and repeat analysis of the compact genome of the

JOURNAL freshwater pufferfish tetraodon nigroviridis
 MEDLINE Genome Res. 10 (7), 939-949 (2000)
 REFERENCE 3 (bases 1 to 1100)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/tetraodon.

FEATURES
 Source location/Qualifiers
 1..1100
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="A"
 /note="Genoscope sequence ID : COA007CD03c1-end : 77"

BASE COUNT 314 a 228 c 252 g 304 t 2 others
 ORIGIN

Query Match 19.3%; Score 177; DB 12; Length 1100;
 Best Local Similarity 55.4%; Pred. No. 4.3e-38;
 Matches 428; Conservative 0; Mismatches 332; Indels 13; Gaps 4;

31 ATTCTGGTGGTGGTGGCCATTTGGGACCTTTGACCTTACATCTGCCAAGCAATG 90
 773 ATCTGTATAGTTCTGCGGTATGACAGAGATTTCATACGACTCTGTAAGTTTAC 714
 91 TTCAACATCCAGAGATCCATCTCACAAGGACCGTTACCTGTACAAACAATCCCATGCA 150
 713 TACAAATTTACAGCGCTGAGCTCAGATGATCGATCCCTTTACGACAAAGCCAGGCT 654
 151 GTTCTGATTCATCCAGAGACATGTTGGATGTGACAAATTTACCTCAGCAAGTAGG 210
 653 GTTCTGTCTTCCACAGAGACATACAGTGGATCTGGAAATCTGCCAGTGGACCCGG 594
 211 CCACCTTTCCAGAAATGATTTGGATGAATTTGGAATCACCACCTCCCAAAAG 270
 593 CCAATTTTTCAGAGGTGATTTGGTTCTACCTGGAATCTCCAGAACACAAATGAGATA 534
 271 AGTGGCATGAGCATTTGTTAACTTACCTGACTGACTTACCGCGCTGATTCAGATATCAA 330
 533 CCAGGCTTGAAACAGATTTCAACATGACTCTGAATCAGAGAAATTCGCAATTTGTA 474
 331 GTGCTTA---TGCTCTTGGAGGTGAAGCAAAATCCCTTGTGTTGAAGTCCAAAGC 387
 473 GCCAGGTACCTCTTACCATCAGAGAGAAATCTGACGGAAGAAATAGTTCTGCCGAG 414
 388 AAAGAGAAATTTGCTGCTGGTGTGAGTAATGAAACCTTGAGC-----ATGCCAGA 441
 413 AAAAACAACCTGCTGCTTGTGATGTGATGATCTACACTGAGGACAGATACAAAGA 354
 442 GTCAAGTATTCATGAGCTGACGAAAGAAATGGAATTCATACCTGAGGCAAGCAATTT 501
 353 GCGCAATTTATATGATCAATCAACATATTAACATTAAGTTGTTTGTGTGTGTAC 294
 502 GGAGAAATATGATCATGATTAATAAATTTGATT---CTTACCAATGCTGTGTAATTTAT 558
 293 ACAGGGTCAAGTTGGAGATTTGATCAATTAATACCCGACCGTGTACCTGTAATTTTAC 234
 559 CTTTCCTTTGAAAATTTCAATCCACAAGATTTACATCAGGAAAGATATCAATAGCTTTT 618
 233 TTATCATTTTGGAACTCATTAACAAGATTTACTACAGAGAGATCAACGACGACCTT 174
 619 CTGGCTGGCTCTGACTGTTGTTGTCGAGACCATCTAGGAAACATATGAAATTAATTT 678
 173 GCACAGAGAACCGTCTTCTGAGTTCTGGGACCCCAAGAGCCACTACGAATGTTCTTC 114
 679 CCACAGATTCATCATTCATGAGGAGATTTATTAATCTCCCACTGAGCTAGCAAAATAT 738
 113 CTTTACAGACTCTTTATTCACGATGATGATTTTGCAGATCCAAAGGACATGGCAGAGCAT 54

QY 739 CTGAGGAGAGTCGACAA-AAACATTAAGTATTAATCTTACTTAACTGGA 790
 Db 53 CTGCTCTCTTGGACAGAGNATGAGAGCATATCATGCGCTACTTATGATGGA 1

RESULT 9
 CDS0429P/c 1016 bp DNA linear GSS 18-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-Or1 end of clone
 DEFINITION 075P02 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION AL271222.1 GI:7993186
 VERSION AL271222.1
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 1016)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

TITLE Unpublished
 JOURNAL 2 (bases 1 to 1016)
 REFERENCE Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Britlier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1016)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/tetraodon.

FEATURES
 Source location/Qualifiers
 1..1016
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="G"
 /note="Genoscope sequence ID : COB075DH01SP1-end :
 PUC-Or1"

BASE COUNT 246 a 216 c 253 g 298 t 3 others
 ORIGIN

Query Match 18.6%; Score 170.2; DB 12; Length 1016;
 Best Local Similarity 57.0%; Pred. No. 3.2e-36;
 Matches 394; Conservative 0; Mismatches 283; Indels 14; Gaps 4;

32 TTTCGGTGTGGGTGGCCATTTGGGACAGCTTTGACCTTACATCTGCCAAGCAATGT 91
 Db 715 TTTCGTGTGGTGTGGCCATTCGGACAGGTTGATTTCAAGAGTTGCTCCACCACAT 656
 QY 92 TCAACATCCAAAGATGATCTCACAACGACGCTGACGTACACAACAAATCCCATGCG 151
 Db 655 TCAAGATTTGACAGCTG-CATCTGACGACGACAGATCTCTGTACAGCAAAAGGAG 597
 QY 152 TTCTGATCCATCCAGGACATGATTGGATGTGACAAATTTTACCTCAGCAAGCTTAGC 211
 Db 596 TCATTTTCTACAAAGACACAACATGAGCACAGACAGCTTCCACAGGCTCCAGGCC 537
 QY 212 CACCTTCGAGAAATGATTTGATGAAATTTGGAATTCACCAATGACATCTCCCAAGA 271
 Db 536 CCATTTCCAGAGGTGGTTGGTTGTTAATGATGATCCCAACAAACATATACGAGAAAC 477

BASE COUNT	199 a	190 c	205 g	245 t	5 others
ORIGIN	/db_xref="taxon:99883" /clone="198K09" /clone_11b="c" probe_Tgenoscope sequence ID : COAG198AF05SP1-end : PUC-011				
Query Match	17.6% Score 161.4; DB 12; Length 844;				
Best Local Similarity	56.2% Pred. No. 8.3e-34;				
Matches 319; Conservative	2; Mismatches 244; Indels 3; Gaps 1;				
OY 25	ACTACTATTCGTGCTGGTGGTGTGGCCATTTGGGACACCTTTGACCTTACATCCTGCCAA 84				
Db 280	ACCACGGGTCGTCCTGCTGGTTCGGCCATTTGGGATTTACCTTTGACCCCAAGTGTGGCAGC 339				
OY 85	GCAATGTCACATCCAAAGATGCGCATTCACCAACGACGACCGCTTACTGTACACAAATTC 144				
Db 340	GCTTACTGTGTGATAGACGGCTGTTTCATTTACAGCGGACAGAACTATTATGAGGAAGA 399				
OY 145	CATGCAAGTTCGATCCATCACCAGACATCATGTTGGGATCTGCACAAATTTACCTCAGCAA 204				
Db 400	ATGCGGATTCGTTATCATCATCAGAGAAATTTCTCCGACCTGTCCCAATCTGCCGACAG 459				
OY 205	GCTAGGCCACCCCTCCAAATAGATTTGGATTTGGATTTGGATTCACCAATTCACACTGCC 264				
Db 460	GAGCGGCGCTTGTACCATATGTTGGTCTGTTCAACATGAGAAATTCCTGTACTCCGAC 519				
OY 265	CAAAAGACGTGCATTTGACACTCTGTTTAACTGACACTGTGATTCACCGCGGATTTGAT 324				
Db 520	AGGTTCTTGGAAATGACAGGCTTGTTCACACTGACTCGAATTTATCGTGTGATGCTGAT 579				
OY 325	ATCCCAAGTCCTTAAAGCTTCTTGACGGTAAACACAAATCCCTTCGTGTTGAAAGTGCA 384				
Db 580	ATTTCATGCTCATGTGTACTGTTCTGTTCTTCAAAAATGAGAGAACTGTGTCCTCC 639				
OY 385	AGCAAAAGAAATTTGTTGCTGCTGGTTTGATTAATGTAACCTTGAGACTGCCAGATC 444				
Db 640	AGCAAAAGATTAAGTGGTCTGCTGATATCTACGCAACCTGGAAGAGACATTAATACGGGTG 699				
OY 445	AAGTTTACCAATGAGCTTAAGCAAAAGAGTGAATTCATACCTTACCGGGACAGATTTGGA 504				
Db 700	AAATTCCTCAAGGAGCTTCCAGACGCAATTAAGATCACCACTTTGACGCTGTTGGG 759				
OY 505	GAATTCATCAATGATTAATTTGATTTCCATCCATATCTGCTTTAAATTTTATCTTCC 564				
Db 760	GGCTA---CAACTGGGGAGATGCTCGTCCATCATGACCGACCTGTAAGTTCCTACCTGCT 816				
OY 565	TTTGAAATTTCAATCCACAAAGATTTACA 592				
Db 817	TTTGAGACTCCATCTACATTTGACTGACATA 844				
RESULT 11	887 bp DNA linear GSS 14-MAY-2000				
LOCUS	Tetradodon nigroviridis genome survey sequence r7 end of clone				
DEFINITION	156120 of library G from Tetradodon nigroviridis, genomic survey				
ACCESSION	AF208693.1 GI:7867512				
VERSION	GSS; genome survey sequence.				
KEYWORDS	Tetradodon nigroviridis.				
SOURCE	Tetradodon nigroviridis				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.				
REFERENCE	1 (bases 1 to 887)				
AUTHORS	Roest-Crolius,H., Jalllon,O., Dasilva,C., Fitzmes,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.				
TITLE	Characterization and repeat analysis of the compact genome of the				

[illegible][illegible]

QY	DB	QY	DB	
431	AATTCAAATTCACGAAAGACTACATCATCTAGAAACCTGGACATAAATGCGCCTCATCTGG	490		
QY	628	TCCTGACCTGTTGTTCTCTGGACCATCTAGGGAAACCTATGAGAATTTATATTCACACAGAT	687	
DB	491	GCTCTACCTGTTGGTTTTGGGGCCCTGGGCTTACTATGACCGTTTTTTTACCCCGCG	550	
QY	668	TCATTCATTCACTGCGAAGATTATTAACCTCTCCACGTGAGCTGACCAAGTATCTGAAGAA	747	
DB	551	TCATTCATCTACAGTGTGATGACCTTCCAACTGCTAAGGGACCTGGCGATTATCTAAACGAG	610	
QY	748	GTCGACAAAACCAATTAAGTTATACCTTGGTTACTTTAACTG	788	
DB	611	CTTGATTAAGATGATTAATTAACATTCGCAACCTTTTCAACTG	651	

RESULT 13	
Bj050300/c	
LOCUS	666 bp mRNA linear EST 06-DEC-2001
DEFINITION	Bj050300 N18B Mochiz normalized Xenopus neurula library xenopus
ACCESSION	laevis cDNA clone X0202015 3', mRNA sequence.
VERSION	Bj050300
KEYWORDS	Bj050300.1 GI:17392988
SOURCE	EST.
ORGANISM	African clawed frog.

ORGANISM
Xenopus laevis; Chordata; Craniata; Euteleostomi;
Euarchyota, Metazoa; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 666)

AUTHORS Y. Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T., and Konata

TITLE Expressed genes in *X. laevis* embryo

JOURNAL Unpublished (2001)

COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsbini@genes.nig.ac.jp.

FEATURES	SOURCE
location/Qualifiers	1. .666
/organism="Xenopus laevis"	
/db_xref="taxon:8355"	
/clone="X1023o15"	
/clone_11b-"NIBB Mochi normalized Xenopus neurula library"	
/tissue_type="whole embryo"	
/dev_stage="stage 15"	
BASE COUNT	178 a 116 c 157 g 215 t
ORIGIN	

Query Match	Similarity	16.6%	Score 151.6	DB 10	Length 666
Best Local	Similarity	34.5%	Fred. No. 4e-31		
Matches 348	Conservative	0	Mismatches 284	Indels 6	Gaps 2
OY	251	CACTGCACACTCCCCAAAAGAGTGGCATTGAGCACTGTTTAACTGACTCTGACTTACC	310		
Db	666	CAGCCCACTATCCCAACACAGCGTTTATGAGAAATACTCATCTGACCACTGCTTTAC	607		
OY	311	GCGGTGATTACATATCCACAGTGCGCTTATGGCTTCTTGACGGTAAACCAATCCCTTGC	370		
Db	606	GTCGTGACTCAATATTTTCTACTCCTTATGGCTGAGTGAAGGCAATGATGAGAAAGAA	547		
OY	371	TGTTTGAAGTGCACAAAGCAAGAAATTGGTGTGCTGGTGTGTGATTAAGTGAACCTTG	430		
Db	546	ACTTCATCTATGTCGGCCCAAGACAAAACATAGTACCTGGTATCTGATTAATGGAACCCCA	487		
OY	431	AGCATGCCAGACTCAAGTATTACAAATGAGCTAAGCAAAAGCAATTGAAATCCATTACCTAG	490		
Db	486	ACTCCAGAAGATCAAAATTTATTTATGAGGAGTTTAAACCCACACCTTCCTGTTGATTAATCTAG	427		

[illegible]

RESULT 14	
CNS034xf/c	
LOCUS	
DEFINITION	CNS034xf
	711 bp DNA linear GSS 15-MAY-2000
	Tetradodon nigroviridis genome survey sequence PUC-ori end of clone
	21,718 of library G from Tetradodon nigroviridis, genomic survey
	sequence.
ACCESSION	AF228012
VERSION	AF228012.1
KEYWORDS	GI:7886987
	GSS: genome survey sequence.

TITLE	AUTHORS	REFERENCE	ORGANISM	SOURCE
Characterization and repeat analysis of the compact genome of the	Weissenbach, J., Bonnau, L., Billault, A., Queller, F., Saurin, W., Bernot, A. and	Roest-Crolius, H., Jallou, O., Dasilva, C., Firames, C., Fisher, C.,	Tetradodon nigroviridis.	Tetradodon nigroviridis.
			Tetradodon nigroviridis	Tetradodon nigroviridis
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
			Acanthomorphae; Acanthopterygii; Percomorphae; Tetradontiformes;	
			Tetradontidae; Tetradodon.	
			1 (bases 1 to 711)	

JOURNAL REFERENCE AUTHORS	freshwater, putrefaction, rhabdovirus, hantaviruses Unpublished 2 (bases 1 to 711) Roest-Croollus, H., Jallion, O., Dasilva, C., Bouneau, T., Fisher, C., Bernot, A., Fitzmes, C., Winker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
TITLE	Human gene number estimate provided by genome wide analysis using Hantavirus nucleocapsid DNA sequence
JOURNAL REFERENCE AUTHORS	Unpublished 3 (bases 1 to 711) Genoscope, Institut National de la Santé et de la Recherche Médicale

FEATURES	SOURCE
<p>LINE</p> <p>JOURNAL</p> <p>COMMENT</p>	<p>Direct submission</p> <p>Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases</p> <p>This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.</p>
	<p>Location/Qualifiers</p> <p>1..711</p> <p>/organism="Tetraodon nigroviridis"</p> <p>/db_xref="taxon:99883"</p> <p>/clone="212578"</p> <p>/clone_1fb="G"</p>

	PUC-Orl"				
BASE COUNT	185	a	153	c	
ORIGIN					
				168	g
				203	t
					2 others

Query Match	16.5%;	Score 151.2;	DB 12;	Length 711;
Best Local Similarity	55.0%;	Pred. No. 5.3e-31;		
Matches 360;	Conservative 1;	Mismatches 284;	Indels 9;	Gaps 3;

Oy 170 ACATCAGTTGGGATGTGCACAAATTTAACTGCACCAAGCTGCGACCCCTTCCAGAAATGGA 229
 Db 705 ACATCAGTGGGAATCTGGCAAACTTCCAGTGGAGCCGGGGCCATTTTTCAGAGATGGG 646
 Oy 230 TTTGATGATTTTGGAAATCCACTCACACTCCCAAAAAGATGGCATTTGAGCACTTGT 289
 Db 645 TTTGGTTCTATGTAAACCTCCAAATGAACACATACAGATTAAGTACTGGTGTAGAAACACTCT 586
 Oy 290 TTAACCTCAGCTGACTACCGCGGTGATTCAGATATCCAGTGGCTTTTGGCTTCTTCA 349
 Db 585 TCAACATGACTCTGTGACTACAGGAAAGACTCCGACATTTGTGGCCAGGTACCTCTTACCA 526
 Oy 350 CGGTAGGACACAATCCCTCTCTGTTTGAAG---TGCCAAAGCAAGAGAAATTTGGTGTCT 406
 Db 525 TCAGAGAGGAAATACTGACGGGAAATATGTTCTGGCGGAGAAAAACAACACTGTGTGT 466
 Oy 407 GGGTTGTGAG---TACTGGAAACCTGAGCATCCAGACTCAAGTATTAACAATGACTTAA 463
 Db 465 GGAATGTCAGCAAAATATGTATGATTCACCTGCTGCTGAGACAAACGAAATTTTCTGTAACATAT 406
 Oy 464 GCMAAGCATTAATATCCATTCCTACCTACGGGCAAGCAT---TTGAGAGAAATATGCAATGATA 520
 Db 405 CCCAGCATATTTACATAAATGTTTATGTGTACGCCCTACACAGGGGTCAAGTTGGAGATTGG 346
 Oy 521 AAAATTTGATTCCTACCATATATCGCTGTGAATTTTATCTTCTTTGAAAATTCATCC 580
 Db 345 ATCAATATTTTCCAGACCGGTGTCTACTGTAAATTTTACTCTGTCCTTTGAGAACTCAATCT 286
 Oy 581 ACAAGGATTAACATACCGGAAAGCTAATACATATGCTTTCTGGCTGGCTCTGTAACTGTGG 640
 Db 285 ACMAAGATTAACATACCGGAAAGGTCACAGACCCCTTCAGTGAAGAACCGTTCCTGTAG 226
 Oy 641 TTCTGGGACCATCTAGGAGAAACTATGAGAAATTAATTCACAGCAATTCATTCATTCAG 700
 Db 225 TTCTGGGACCCCAAGGACCACTAGGAATGTCTTCCCTTCAGACTCTTATTATACA 166
 Oy 701 TGGAAATATTATACCTCTCCAGTGAAGCTAGCAAAAGTATCTGAAAGAGTGCACAAAACA 760
 Db 165 TGAATGATTTTCCAGATCCAAAGGCACTGGCAGAGCATCTGCTCTCTTGACAAAGGATG 106
 Oy 761 AATAATTAATACCTAGTACTCTTAATCTGAGAGAAAGATTCCTACTGTAATATCTTC 814
 Db 105 AGGAGCATACATGGCCTACTTGTAGTGGAGGAAGTACTTACGGCATATCTTCCTC 52

RESULT	15
B41073/c	
LOCUS	
DEFINITION	501 bp DNA linear GSS 18-OCT-199
ACCESSION	HS-1055-B3-G01-ME.abi CIT Human Genomic Sperm Library c Homo
VERSION	sapiens genomic clone Plate=CT 774 Col=2 Row=N, DNA sequence.
	B41073
	B41073.1 GI:2545325

REFERENCE	1 (bases 1 to 501)
AUTHORS	Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S., Traloff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE	Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors
JOURNAL	Unpublished (1997)
COMMENT	Contact: Mahairas GG, Zackrone KD, Hood L

University of Washington
Seattle, WA 98195, USA
Tel.: (206) 616-8744
Fax: (206) 685-7301
Email: krackron@u.washington.edu
Sequence Tagged Connector
Plate: CT 774 row: N column: 2
Class: BAC ends
High quality sequence stop: 501.
location/Qualifiers
1. 501

BASE COUNT ORIGIN	179 a	95 c	83 g	143 t	1 others
----------------------	-------	------	------	-------	----------

Query Match	16.5%	Score 151;	DB 12;	Length 501;
Best Local Similarity	100.0%	Pred. No. 5.4e-31;		
Matches 151; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	GTGTAACTTAACTTAACTGAGAGAGAGATTCACTGTAAATCTCCACAGTTTGG	824
Db	GTGTAACTTGTACTTTAACTGGAGAGAGATTCACTGTAAATCTCCACAGTTTGG	442
QY	GGATACACATCATGTTGGCTGGCGATCATGTGAAAAGCATCAAGATATTAAGTCGT	884
Db	GGATACACATCATGTTGGCTGGCGATCATGTGAAAAGCATCAAGATATTAAGTCGT	382
QY	TGCTAATTAGAAATGCTTTTGGAAATTAA	915
Db	TGCTAATTAGAAATGCTTTTGGAAATTAA	351

Search completed: October 6, 2002, 07:16:27
Job time : 773.105 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 22:42:00 ; Search time 18.6694 Seconds
(without alignments)
12038.636 Million cell updates/sec

Title: US-09-744-748-5_COPY_460_1374

Perfect score: 915
Sequence: 1 accaaactgattatttaa.....agaatggtttggaattaa 915

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/1na/5a_COMB.seq:*
2: /cgn2_6/prodata/2/1na/5b_COMB.seq:*
3: /cgn2_6/prodata/2/1na/6a_COMB.seq:*
4: /cgn2_6/prodata/2/1na/6b_COMB.seq:*
5: /cgn2_6/prodata/2/1na/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/1na/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	13.9	1814	2 US-08-483-151-1	Sequence 1, Appl1
2	127	13.9	1814	2 PCT-US96-06427-1	Sequence 1, Appl1
3	113.4	12.4	1654	1 US-07-914-281-13	Sequence 13, Appl1
4	113.4	12.4	1654	1 US-08-393-246-13	Sequence 13, Appl1
5	113.4	12.4	1654	1 US-08-525-058A-13	Sequence 13, Appl1
6	113.4	12.4	1654	2 US-08-696-731-13	Sequence 13, Appl1
7	113.4	12.4	1654	4 US-09-042-531-13	Sequence 13, Appl1
8	112.2	12.3	1086	1 US-07-914-281-12	Sequence 12, Appl1
9	112.2	12.3	1086	1 US-08-393-246-12	Sequence 12, Appl1
10	112.2	12.3	1086	1 US-08-525-058A-12	Sequence 12, Appl1
11	112.2	12.3	1086	2 US-08-696-731-12	Sequence 12, Appl1
12	112.2	12.3	1086	4 US-09-042-531-12	Sequence 12, Appl1
13	112.2	12.3	2042	3 US-09-063-237-2	Sequence 2, Appl1
14	112.2	12.3	2043	1 US-07-914-281-1	Sequence 1, Appl1
15	112.2	12.3	2043	1 US-08-393-246-1	Sequence 1, Appl1
16	112.2	12.3	2043	1 US-08-273-411-4	Sequence 4, Appl1
17	112.2	12.3	2043	1 US-08-525-058A-1	Sequence 1, Appl1
18	112.2	12.3	2043	2 US-08-696-731-1	Sequence 1, Appl1
19	112.2	12.3	2043	4 US-09-042-531-1	Sequence 1, Appl1
20	112.2	12.3	2043	5 PCT-US91-00899-1	Sequence 1, Appl1
21	111.8	12.2	1316	1 US-07-914-281-10	Sequence 10, Appl1
22	111.8	12.2	1316	1 US-08-393-246-10	Sequence 10, Appl1
23	111.8	12.2	1316	1 US-08-525-058A-10	Sequence 10, Appl1
24	111.8	12.2	1316	2 US-08-696-731-10	Sequence 10, Appl1
25	111.8	12.2	1316	4 US-09-042-531-10	Sequence 10, Appl1
26	100.8	11.0	1068	4 PCT-US91-00899-13	Sequence 13, Appl1
27	100.8	11.0	1215	5 PCT-US91-00899-5	Sequence 5, Appl1

28	100.8	11.0	1488	1 US-07-914-281-9	Sequence 9, Appl1
29	100.8	11.0	1488	1 US-08-393-246-9	Sequence 9, Appl1
30	100.8	11.0	1488	1 US-08-525-058A-9	Sequence 9, Appl1
31	100.8	11.0	1488	2 US-08-696-731-9	Sequence 9, Appl1
32	100.8	11.0	1488	4 US-09-042-531-9	Sequence 9, Appl1
33	100.8	11.0	2175	4 US-08-482-073-9	Sequence 9, Appl1
34	100.8	11.0	2861	4 US-08-482-073-10	Sequence 10, Appl1
35	100.8	11.0	3647	1 US-07-914-281-7	Sequence 7, Appl1
36	100.8	11.0	3647	1 US-08-393-246-7	Sequence 7, Appl1
37	100.8	11.0	3647	1 US-08-525-058A-7	Sequence 7, Appl1
38	100.8	11.0	3647	2 US-08-696-731-7	Sequence 7, Appl1
39	100.8	11.0	3647	4 US-09-042-531-7	Sequence 7, Appl1
40	100.8	11.0	3647	5 PCT-US91-00899-4	Sequence 4, Appl1
41	97.6	10.7	2134	2 US-08-483-151-3	Sequence 3, Appl1
42	97.6	10.7	2134	5 PCT-US96-06427-3	Sequence 3, Appl1
c 43	37.2	4.1	1728	1 US-08-403-866-12	Sequence 12, Appl1
c 44	37.2	4.1	2231	1 US-08-403-866-14	Sequence 14, Appl1
c 45	37.2	4.1	12720	1 US-08-403-866-11	Sequence 11, Appl1

ALIGNMENTS

```
RESULT 1
; Sequence 1, Application US/08483151
; Patent No. 5858752
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Holgersson, Jan
; TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,151
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/278001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1814 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-483-151-1
```

Query Match 13.9%; Score 127; DB 2; Length 1814;

Best Local Similarity 52.3%; Pred. No. 9,7e-29;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

```
Oy 104 GATGCATCTTACACAGGACGCTGCTACACCAATCCCATGCGATTGATTCATC 163
Db 548 GCTGCGCTCTAGTGTCTTACCGGAGACCTGCTAGCCAGTGTGATGCTTCCACC 607
```

```
QY 164 ACCGAGATCATCACTGGATGCTGACAAATTTACCTGACGAGCTAGGCCACCTTCAG 223
DB 608 ACCGTAGCTGCAAAACCCGCAATCTCTCCATCCCTGGACGACGAGCCACGAGCAGC 667
QY 224 AATGATTTGATGATTTTGGAAATTCACCACTACACATCCCAAAAAGATGATGAGC 283
DB 668 CTTGGGCTGGGCTCCATTCAGATTCGCCCCAGTAATCCATGCTTCACGCTTCGGG 727
QY 284 ACTGTTTACCTGATCTGATCTACCGGCTGATTCAGATTCAGATTCAGATTCAGAT 343
DB 728 GCATCTTCACTGGGTGCTAGCTATCGGCTGATTCAGATTCAGATTCAGATTCAGAT 787
QY 344 TCTTGACGGTAAACGACAAATCCCTGCTGTTTGAAGTCCCAAGAAAGAAATTTGGT 403
DB 788 GCTTGAGAGCTCTCTCTGGGCC---CACATCCCACTACCGGCAAAAGACGATGCTG 844
QY 404 GCTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 463
DB 845 CCGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904
QY 464 GCAAAAGCATTTGAATTCATCTACGAGGAGCATTTGGAGAAATGTCATGATGATA 523
DB 905 CCGCTCATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
QY 524 ATTGATTCCTACCATATCTGCTTGAATTTTATCTTCTTGAATTTCAATTCACA 583
DB 965 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
QY 584 AGGATTACATCAAGGAAAGCTAT--ACAATGCTTTTCTGCTGCTGCTGCTGCTGCTG 640
DB 1025 GGGACTACATCACTGAGAAATTTCTGCGGCAATCCCTGCGGCTGCTGCTGCTGCTG 1084
QY 641 TCTTGAGGACCATCTAGGAAATCTATGAGAAATTAATTCAGAGATTTCTTATTCATG 700
DB 1085 CGCTGGACCTCTCTGCGGCACTACGAGGCTTTTGGCCACAGATGCTTGTACAG 1144
QY 701 TGGAGATTAATCACTCCAGAGTACGAGTACG 731
DB 1145 TGGAGCATTCAGCTCTGCGGCTGAGACTGCG 1175
```

RESULT 2

Sequence 1, Application PC/TUS9606427

GENERAL INFORMATION:

APPLICANT: The General Hospital Corporation

TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/06427

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/483,151

FILING DATE: 07-JUN-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Leach, Karen F.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/278W01

TELECOMMUNICATION INFORMATION:

```
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1814 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US96-06427-1
```

Query Match 13.9%; Score 127; DB 5; Length 1814;
Best Local Similarity 52.3%; Pred. No. 9, 7e-29;
Matches 330; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

```
QY 104 GATGCTATCTACACGAGACCGTTCACTGATACAAATTCATGCAAGTTCATTCATC 163
DB 548 GCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
QY 164 ACCGAGATCATCACTGGATGCTGACAAATTTACCTGACGAGCTAGGCCACCTTCAG 223
DB 608 ACCGTAGCTGCAAAACCCGCAATCTCTCCATCCCTGGACGACGAGCCACGAGCAGC 667
QY 224 AATGATTTGATGATTTTGGAAATTCACCACTACACATCCCAAAAAGATGATGAGC 283
DB 668 CTTGGGCTGGGCTCCATTCAGATTCGCCCCAGTAATCCATGCTTCACGCTTCGGG 727
QY 284 ACTGTTTACCTGATCTGATCTACCGGCTGATTCAGATTCAGATTCAGATTCAGATTC 343
DB 728 GCATCTTCACTGGGTGCTAGCTATCGGCTGATTCAGATTCAGATTCAGATTCAGATTC 787
QY 344 TCTTGACGGTAAACGACAAATCCCTGCTGTTTGAAGTCCCAAGAAAGAAATTTGGT 403
DB 788 GCTTGAGAGCTCTCTCTGGGCC---CACATCCCACTACCGGCAAAAGACGATGCTG 844
QY 404 GCTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 463
DB 845 CCGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904
QY 464 GCAAAAGCATTTGAATTCATCTACGAGGAGCATTTGGAGAAATGTCATGATGATA 523
DB 905 CCGCTCATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
QY 524 ATTGATTCCTACCATATCTGCTTGAATTTTATCTTCTTGAATTTCAATTCACA 583
DB 965 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
QY 584 AGGATTACATCAAGGAAAGCTAT--ACAATGCTTTTCTGCTGCTGCTGCTGCTGCTG 640
DB 1025 GGGACTACATCACTGAGAAATTTCTGCGGCAATCCCTGCGGCTGCTGCTGCTGCTG 1084
QY 641 TCTTGAGGACCATCTAGGAAATCTATGAGAAATTAATTCAGAGATTTCTTATTCATG 700
DB 1085 CGCTGGACCTCTCTGCGGCACTACGAGGCTTTTGGCCACAGATGCTTGTACAG 1144
QY 701 TGGAGATTAATCACTCCAGAGTACGAGTACG 731
DB 1145 TGGAGCATTCAGCTCTGCGGCTGAGACTGCG 1175
```

RESULT 3

US-07-914-281-13

Sequence 13, Application US/07914281

Patent No. 5324663

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

[illegible]

RESULT 5
 US-08-525-058A-13
 Sequence 13, Application US/08525058A
 Patent No. 5770420
 GENERAL INFORMATION:
 APPLICANT: LOWE, JOHN B.
 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
 STREET: 1755 Jefferson Davis Highway, Fourth Floor

City: Arlington
 State: Virginia
 Country: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525,058A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavallee, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1654 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-525-058A-13

Query Match	12.4%	Score 113.4	DB 1	Length 1654
Best Local Similarity		Pred. No. 1.2e-24		
Matches 401	Conservative	0	Mismatches 361	Indels 21
			Gaps	
OY	ATATCGGTGGGGTGGGTCATTTTGGGCGAGACCTTGTGACCTTATACATGCTCGCAAGCAATG	90		
Db	ATCTCGGTGGGACGCGGCTTTTAAACAAACCATAGCTCTGCCCCGCTGCTAGAGATG	377		
OY	91 TTT----CAACATCCAGATATGCCATCTCACACGACCGCTTACGTATACAAACAAATCCCAT	147		
Db	GTGCGCTGGCAGCGCTGACTGACATCATCTGCGACCCCAAGGTGTATCCACAGCGAC	437		
OY	148 GCAGTTCTGATTCATACCGGACGACATCACTTGGATGTGA--CAAAATTACCTCAGCAA	204		
Db	438 GCGGTCTATCTGTGACACCGAGGTCATGTCAACCCCACTGCGCAGTCCACAGCTCC	497		
OY	205 GCTATGCCACCCCTTCAGAAATGGATTGGATGGAAATTTGGATTGATACCACTGTCACTGCC	264		
Db	498 CCGAGGGGACAGGGGACGAGTATGATCTGTTCACTATGAGTCCCAAGCAGCTGTGG	557		
OY	265 CAAAGAGTGGACATTGAGCACTGTTTAACTGCATCTGCATCTTACCGCGGATTCAGAT	324		
Db	558 CAGGTGAAGAACATGGACGGAATCTTCATCTCACATATGCTCTTACCGAGAGCATCCGAC	617		
OY	325 ATCCAACTGCTTATGGCT-----TCTTGACGGTAAAGCAAAATCCCTTCGTGTTT	375		
Db	618 ATCTTCACGGCTTACGCTGTGCTGACGCCGTGCTCGGACAGCCTGCCACCCACGCTGC	677		
OY	376 GAATGGCCAAAGCAAGAAATTTGTTGCTGGTGTGTGAGATTAACCTGGAAACCTTGAGAT	435		
Db	678 AACCTCTGGCCCAAGAACCGAGAGCTGGGTGGGACAGTGTCCAATGGGGGCCAAATCTCC	737		
OY	436 GCCAGATCAAGATTAATCAATGAGTCAATACCAAAACCAATTTGAAATCCATACCTACGAGCAA	495		
Db	738 GCCAGGCTGCGTACTACCGACCGCTGACGGCCCATCTCAAGGTGACGTGTACGG---A	794		
OY	496 GCATTGGAGAAATATGTCATGATTAATAAATTTGATTCCTACCATATCTGCTTGAATTT	555		
Db	795 CGCTCCCAAGCCCTCGCCCCGAGGAACCATGATGAGAGCGTGTCCCGGTACAAATTC	854		
OY	556 TATCTTCTTTTGAATTAATTCACCAAGATTAATCATCAGGAAAAAGCTT---ACAAT	612		
Db	855 TATCTGCTTCGAGACTCTCTTGACACCCGACTATATATCCGAAACATCTGTGAGGAGAC	914		

```

Oy      613 GCTTTTCGGGCGTCTGTCTGTTTGCTGGAGCATCTAGGGAATAACTATGAGAAAT 672
Db      915 GCCCTGGAGAGCCCTGGGCCGTGCCCTGCTGTGGGCCCCACACAGACATCAGAGAGG 974
Oy      673 TATATTCCAGCAGATTCATTCATTCATGTGGAGAATTAATTAATCTCCACGTAGAGTAGCA 732
Db      975 TTCCGTCAGACCAGACGCCTTCATTCACAGGTGGACGACCTTCAGAGAGCCCCAAGAGACTGGCC 1034
Oy      733 AAGTATCTCGAAGGAAGTGCACAAAAACAATTAAGTTATACCTTAGTACTTAACTGAGAG 792
Db      1035 CGGTACCTCGAGAGAGCTGGACACAGACACCGCCGCTACCTAGACTACTTCGCTGGCGG 1094
Oy      793 AAG 795
Db      1095 GAG 1097

RESULT 6
US-08-696-731-13
; Sequence 13, Application US/08696731
; Patent No. 5955347
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US//08/696,731
; FILING DATE: 14-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/393,246
; FILING DATE:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-696-731-13

Query Match          12.4%; Score 113.4; DB 2; Length 1654;
Best Local Similarity 51.2%; Pred.No.1.2e-24;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
```

31 ATTGGGTGGTGGTGGCCATTGGGAGACCTTTGACCTTACATCTGTCCAGCAATG 90
 318 ATCTGCTGTGAGCTGGCTTTTAAACAACCAATAGCTGCCCCGCTGCTCGAGAGT 377
 91 TT---CAACATCCAAAGATGCGCATCTCTACAACGAGCGCTTCACTGTACCAAAATCCAT 147
 378 GTGCTTGCCACGGCTGATGTGCAACATCTACTGCGACCGCAAGGTGTATCCAGGCAAC 437
 148 GCAGTTCTGATTCATCACCGAGACATCATGTTGGATCTGA---CAATTTACTTAGCA 204
 438 GCGGCTCATCTGTCAACACCGAGAGGCTATGTAAACACCCAGTGCACCTCCACAGCTCC 497
 205 GCTAGCCACCGCTTCCAAATATGATTTGGATTTGGAAATTCACCAATCTACACTCCC 264
 498 CCGAGGCGGAGGGGACCGATGATCTGGTTTACCAATGAGAGTCCCAAGGCTACTGCTGG 557
 265 CAAAGAGTGGCATTTGACACATCTGTTTAACCTGACTGTGATTAACCGCGGATTCAGAT 324
 558 CAGCTGAAGAGCCATGAGAGGATCTTCAATCTCAACATGTCTTACCGGAGGACTCCGAC 617
 325 ATCCAGTGCCTTATGCT-----TCTTGACGCTAAGCACCAATCCCTTGCTGTTT 375
 618 ATCTTCACGCGCTTACGGCTGCTGAGCGCGTGTCCGCGCAGCTGCCACCCACCGCTC 677
 376 GAAGTGCACAAAGAAATATGTTGTCGTGGTGTGGATTAATCTGGAACCCCTGACAT 435
 678 AACCTTGGCGCCACGACCGAGCTGGTGGCTGGGCAAGTGTCCAACTTGGGGGSCAAATCTCC 737
 436 GCCAGAGTCAAGTATTACATGAGCTTAAGCAAAAGCATTTGAAATCTCATCTACGGGCAA 495
 738 GCCAGGGTGGCGTACTACAGAGCTGCAAGGCCATCTCAAGTGGAGCTGTACGG---A 794
 496 GCATTGGAGATATGTCAATGATTAATAATTTGATTCCTACCATATCTGCTTGTAAATTT 555
 795 CGCTCCCAAGCCCTCCGCCAGGAAACATGATGAGAGCGCTGTCCCGGTACAAAGTTC 854
 556 TATCTTCTTGAATAATTCATCCACAAGATTATCATACGGAATAACCTATP---ACAT 612
 855 TATCTGGCTTCTGAGAACTCTTGTGACCCCGATCTACATACCGAGAACTCTGGAGGAC 914
 613 GCTTTTCTGGCTGCTGTACTCTGTGTTGTTGAGGACCATTAGGAAACTATGAGAT 672
 915 GCCCGTAGGCGCTGGGCGCTGCGCTGGTGGTGGGCCCCACAGAAAGCAACTAGAGAGG 974
 673 TATATTCAGAGATTCATTATCTATGTGGAAGATTAACTCTCCAGTAGAGTAGCA 732
 975 TTCCGCGCACCCGAGCGCTTATCATCGCTGGACGACTTCCAGAGCCCAAGGACTGTGACC 1034
 733 AAGTATCGAAGGAAGTGCACAAAACAATTAAGTATACCTAGTTACTTTTAACGAGAG 792
 1035 CGGTACTCTGCAAGAGACTGACAAAGAACAGCGCGCTACTGTAGCTACTTGTGCTGGCG 1094
 793 AAG 795
 1095 GAG 1097

RESULT 7
 US-09-042-531-13
 : Sequence 13, Application us/09042531
 : Patent No. 6268193
 : GENERAL INFORMATION:
 : APPLICANT: LOWE, JOHN B.
 : TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 : OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 : TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 : OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
 : ADDRESSEE: P.C.
 : STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/042,531
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1654 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-13

Query Match 12.4%; Score 113.4; DB 4; Length 1654;
Best Local Similarity 51.2%; Pred. No. 1.2e-24;
Matches 401; Conservative 0; Mismatches 361; Indels 21; Gaps 5;
QY 31 ATTCTGTGGTGGTGGCCATTGGGCGAGACCTTTGACCTTACATCTGCGCAACCAATG 90
DB 318 ATCTGTGTGGTGGTGGCCATTGGGCGAGACCTTTGACCTTACATCTGCGCAACCAATG 377
QY 91 TT--CAATCCAGATGCGATCTCTACACGACGCTTCACTGTACACCAATCCCAT 147
DB 378 GTGCTGTGGTGGTGGCCATTGGGCGAGACCTTTGACCTTACATCTGCGCAACCAATG 437
QY 148 GCAGTTTGCATCCATCACCAGACATCAGTTGGATCTGA--CAATTTTACTCAGCAA 204
DB 438 GGGGTATCGTGCACACCAAGAGTGTATGTCACACCCAGTGGCCAGCTCCAGCTCC 497
QY 205 GCTAGGCCACCTTCCAGAAATGATTTGATGATTTGATTTGATTTGATTTGATTTGATTTG 264
DB 498 CCGAGGCGGCGAGGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 557
QY 265 CAAAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 324
DB 558 CAGCTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 617
QY 325 ATCCAGTCCCTTATGGCT-----TCTTGAAGGTAAAGCACAATTCCTTCTGTTT 375
DB 618 ATCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTAC 677
QY 376 GAATGCGCAAGAGAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 435
DB 678 AACCTTGGCGCAAGACCGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 737
QY 436 GCCAGAGTCAAGTATTAATGAGTAAAGCAAAAGCATTTGAATTCATACCTACGAGGCA 495
DB 738 GCCAGGTTGCTTACTACAGAGCTGAGGCCCATTCATGAGTGTGAGTGTGAGTGTGAGT 794

QY 496 GCATTTGGAGATATGTCATGATATAAATTTGATTTCTTACCATATCTGCTGTAATTT 555
DB 795 CGCTCCACACMACCCCTGCCCCAGGAGAACCATGATGAGAGAGCTGTCCCGGTACAGTTTC 854
QY 556 TATCTTTCTTTGAAATTCATTCACAGATTTATCATCAGGAAAGCTAT--ACAT 612
DB 855 TATCTGCTTGTGAGATCTCTGAGTCCCGACTTACATCCAGAGAACCTTGTGAGGAGAC 914
QY 613 GCTTTTGTGCTGCTGTGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 672
DB 915 GCCCTGAGGCTGTGGCCGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 974
QY 673 TATATTCAGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 732
DB 975 TTCTGCGCACCGACGCTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1034
QY 733 AAGTATCTGAAGAGTGCAGCAAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 792
DB 1035 CGGTACTGTGACAGCTGTGACAGAGCACCGCCCTACTGAGTACTGCTGCGCG 1094
QY 793 AAG 795
DB 1095 GAG 1097

RESULT 8
US-07-914-281-12
Sequence 12 Application US/07914281
Patent No 5324653

GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOPOLYDIPS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, WAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-914-281-12

Query Match 12.3%; Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

```

QY 106 TGCATCTCAACGACGCGTTCATGACAAATCCCATCAGTTCATCATCAC 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TGCACATCATCTCCGACGCGCAAGGTGTACCCACAGCAGACACGCTCATCTGCACAC 330
QY 166 CGAGACATCATGTTGGG---TTCGACAAATTTTCTCAGACGACCTTCCAG 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TGGGATATCATGTTCCAAACCTTAAGTACGCGCTCCACCTTCCCGAGGCGCGAGGGG 390
QY 223 AATGATTTGATGATTTGATGATTCACCAATCAGCTACCTCCCAAAAGAGTGGCATTTAG 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CGCTGATCTGTTCAACTTGGAGCCACCCCTTAAGTACGAGACCTGGAAGCCCTGGAC 450
QY 283 CACTTGTTAACCTGACTGACTGACTACCGCGGTGATTCAGATATCCAGTCCCTATGCG 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AGATATCTCAATCTCAACATGCTCTACCGGAGGAGCTCCGACATCTTCACGCGCTTAC 510
QY 343 TTCTTGACGTTAAGCAAAATCCCTTCGT-----GTTTGAATGCCAACAAGAG 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 TGGCTGAGCGGTGTGTCGGCGCCGCTGCCACCCAGCTCAACCTCTCGGCGCAAGAC 570
QY 394 AATTTGTTGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
QY 454 AATGACCTAAGCAAAAGCATTTGAATCCATCTACGCGGCAACGATTTGAGAAATATGTC 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAGAGCTCAGAGCTCATCTCAAGGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 690
QY 514 AATGATTAATAATTTGATTTCTCATATCTGCTTGTAAATTTTATCTTCTTGTAAAT 573
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 AAGGGAGCA--TGATGAGAGCGCTGTCCGCTTACAAATCTTACCTGCGCTTGTGAGAC 747
QY 574 TCAATCCACAAGATTTACATACGAGAAAGCTAT---ACAATGCTTTTCTGCGTGTGCT 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 TCTTGTGACCGCCGCTACATCATACGAGAGAGCTGTGAGAGAGCCCTGAGGCTTGGCC 807
QY 631 GTACCTGTGTTCTGAGGACATCTAGGAGAAACTATGAGATTTATTTCCAGAGATTA 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
QY 691 TTATTCATGTGAGAAATTTATCTCTCCAGAGAGCTAGCAAAAGTATCTGAAAGAGTC 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 TTATTCATGTGAGAAATTTATCTCTCCAGAGAGCTAGCAAAAGTATCTGAAAGAGTC 927
QY 751 GACAAAACAAATTAATTTATCTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 GACAAAGACACGCGCGCTTACTGAGCTTCTGCTGCGGAGAG 972

```

```

; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavelley, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1086 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-393-246-12

Query Match 12.3%; Score 112.2; DB 1; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 106 TGCATCTCAACGACGCGTTCATGACAAATCCCATCAGTTCATCATCAC 165
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TGCACATCATCTCCGACGCGCAAGGTGTACCCACAGCAGACACGCTCATCTGCACAC 330
QY 166 CGAGACATCATGTTGGG---TTCGACAAATTTTCTCAGACGACCTTCCAG 222
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TGGGATATCATGTTCCAAACCTTAAGTACGCGCTCCACCTTCCCGAGGCGCGAGGGG 390
QY 223 AATGATTTGATGATTTGATGATTCACCAATCAGCTACCTCCCAAAAGAGTGGCATTTAG 282
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 CGCTGATCTGTTCAACTTGGAGCCACCCCTTAAGTACGAGACCTGGAAGCCCTGGAC 450
QY 283 CACTTGTTAACCTGACTGACTGACTACCGCGGTGATTCAGATATCCAGTCCCTATGCG 342
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AGATATCTCAATCTCAACATGCTCTACCGGAGGAGCTCCGACATCTTCACGCGCTTAC 510
QY 343 TTCTTGACGTTAAGCAAAATCCCTTCGT-----GTTTGAATGCCAACAAGAG 393
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 TGGCTGAGCGGTGTGTCGGCGCCGCTGCCACCCAGCTCAACCTCTCGGCGCAAGAC 570
QY 394 AATTTGTTGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
QY 454 AATGACCTAAGCAAAAGCATTTGAATCCATCTACGCGGCAACGATTTGAGAAATATGTC 513
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CAGAGCTCAGAGCTCATCTCAAGGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 690
QY 514 AATGATTAATAATTTGATTTCTCATATCTGCTTGTAAATTTTATCTTCTTGTAAAT 573
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 AAGGGAGCA--TGATGAGAGCGCTGTCCGCTTACAAATCTTACCTGCGCTTGTGAGAC 747
QY 574 TCAATCCACAAGATTTACATACGAGAAAGCTAT---ACAATGCTTTTCTGCGTGTGCT 630
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 TCTTGTGACCGCCGCTACATCATACGAGAGAGCTGTGAGAGAGCCCTGAGGCTTGGCC 807
QY 631 GTACCTGTGTTCTGAGGACATCTAGGAGAAACTATGAGATTTATTTCCAGAGATTA 690
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
QY 691 TTATTCATGTGAGAAATTTATCTCTCCAGAGAGCTAGCAAAAGTATCTGAAAGAGTC 750
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-696-731-12

Query Match 12.3%; Score 112.2; DB 2; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 106 TGGCATCTCAAAAGCAGCTTACTGATCAACAATCCATGATGCTGATCATCAC 165
DB 271 TGGCAGATCTCAAGCTGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 330
QY 166 CGAGACATCTAGTGGG---TCTGACAAATTTACCTGAGCAGAGTACGACCTTCAG 222
DB 331 TGGGATATCTAGTCCAAAGCTTAAAGTACGCTCCACCTTCCCGAGGCGCAGG 390
QY 223 AAATGGATTTGATGATTTGGAATTCACCACTACACTCCCAAAAGAGTGGCATGAG 282
DB 391 CGCTGGATCTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAGGAG 450
QY 283 CACTTGTCTTAACTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAG 342
DB 451 AGATATCTTCAATCTCAAGCTTAAAGTACGCTCCAGAGGAGGAGGAGGAGGAG 510
QY 343 TTCTGACGCTAGACAAATTCCTTCTG---GTTGAAAGTCCAGCAAGAAAGAG 393
DB 511 TGGCTGAGGCGGTGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 570
QY 394 AAATGGTGTGCTGATGATTTGGAATTCACCACTACACTCCCAAAAGAGTGGCATGAG 453
DB 571 GAGCTGTGCTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAGGAG 630
QY 454 AATGAGCTTAAGCAAAAGCATTTGAATTCACCACTACACTCCCGAGGAGGAGGAG 513
DB 631 CAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 690
QY 514 AATGATTAATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 573
DB 691 AAGGGGAGCA---TGAAGGAGAGGCTGCTCCGAGTCAAGTTTCAAGCTTCAAG 747
QY 574 TCAATCCACAGAGTATTCACGAGAAAGCTAT---ACATGCTTTTGGCTGCTGCT 630
DB 748 TCCTTGACCCCGACATTCATCAGGAGAACTGTGAGAGAACTGTGAGAGGCTGAG 807
QY 631 GATACCTGTGCTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAG 690
DB 808 GTGCTGCTGCTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAG 867
QY 691 TTGATTCATGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAGGAG 750
DB 868 TTGATTCATGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAGGAG 927
QY 751 GACAAAAAACAATTAATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTG 795
DB 928 GACAAAGACAGGCGGCTACCTGAGCTTCTGCTGAGGAGGAG 972

RESULT 12
US-09-531-12
; Sequence 12, Application us/09042531
; Patent No. 6268193
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,531
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavelle, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-042-531-12

Query Match 12.3%; Score 112.2; DB 4; Length 1086;
Best Local Similarity 51.6%; Pred. No. 2.2e-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

QY 106 TGGCATCTCAAAAGCAGCTTACTGATCAACAATCCATGATGCTGATCATCAC 165
DB 271 TGGCAGATCTCAAGCTGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 330
QY 166 CGAGACATCTAGTGGG---TCTGACAAATTTACCTGAGCAGAGTACGACCTTCAG 222
DB 331 TGGGATATCTAGTCCAAAGCTTAAAGTACGCTCCACCTTCCCGAGGCGCAGG 390
QY 223 AAATGGATTTGATGATTTGGAATTCACCACTACACTCCCAAAAGAGTGGCATGAG 282
DB 391 CGCTGGATCTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAGGAG 450
QY 283 CACTTGTCTTAACTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAG 342
DB 451 AGATATCTTCAATCTCAAGCTTAAAGTACGCTCCAGAGGAGGAGGAGGAGGAG 510
QY 343 TTCTGACGCTAGACAAATTCCTTCTG---GTTGAAAGTCCAGCAAGAAAGAG 393
DB 511 TGGCTGAGGCGGTGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 570
QY 394 AAATGGTGTGCTGATGATTTGGAATTCACCACTACACTCCCAAAAGAGTGGCATGAG 453
DB 571 GAGCTGTGCTGATGATTTGGAATTCACCACTACACTCCCGAGGAGGAGGAGGAG 630
QY 454 AATGAGCTTAAGCAAAAGCATTTGAATTCACCACTACACTCCCGAGGAGGAGGAG 513
DB 631 CAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 690

OY 514 AATGATMAAATTTGATCTACATATCTGCTTGAATTTATCTTCTTGAAT 573
DB 691 AAGGGACCA---TGATGAGACGCTGCTCCGGTACAGATCTTACCTGCTTGAAG 747
OY 574 TCAATCCACAGATATACATACGGAAGAACTAT---ACAATCTTTCTGCTGCTCT 630
DB 748 TCTTGTACCCCGACTACATACCGAAGCTGTGAGAGACCCCTGAGGCTGAGCC 807
OY 631 GACCTGTGTCTGTGACCATCTAGGAAACTATGAAATTAATTCAGCAGATTCA 690
DB 808 GTGCCCCGTGTGTGAGCCCGCCAGAGAGACACTACAGAGCTTCCGACCCGAGCC 867
OY 691 TTCAATCATGTGAGAGATTAATACCTCCAGTACAGTACAGAAATATCTGAAGAGTC 750
DB 868 TTCAATCATGTGAGAGACTTCCAGAGCCCAAGACCTGCGCTGACCTGAGAGCTG 927
OY 751 GACAAACAAATAAGTTATACCTTACTTACTTACTGAGAGAG 795
DB 928 GACAAAGACCAAGCCCGCTACCTGAGACTTTCGCTGCGGAG 972

RESULT 13

US-09-063-237-2
Sequence 2, Application US/09063237
Patent No. 6124267
GENERAL INFORMATION:
APPLICANT: McEver, Rodger P.
APPLICANT: Cummings, Richard D.
TITLE OF INVENTION: O-glycan Inhibitors of Selectin Mediated
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Patrea L. Padst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30306-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/063,237
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/649,802
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Padst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF110CIP7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-063-237-2

Query Match 12.3%; Score 112.2; DB 3; Length 2042;
Best Local Similarity 31.6%; Pred. No. 36-24;
Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;
OY 106 TGCCATCTCAACAGGACCGTTCATGTACAAACAAATCCATGCACTTGTGATCATCAC 165

DB 343 TGCCATCATACAGCCGACCGCAAGGTATGCCACAGGCAAGCATGCTATGACCCAG 402
OY 166 CGAGACATAGTTGGA---TCTGACAAATTAATCTACAGCAAGCTAGCCACCTTCAG 222
DB 403 TGGATATATATGTCACACCTTAAGTACAGCTCCACCTTCCCGGCGGAGGAGC 462
OY 223 AATGATTTGATGAATTTGGAATACCACTACACACTCCCAAGAGAGTGCATTAG 282
DB 463 CGCTGATGTGTTCAACTTGGAGCCACCCCTTAATCTGCACACCTGGAACCCCTGAG 522
OY 283 CACTTGTTAACCTGACTGTGACTTACCCCGGATTCAGATATCCAAATGCTTATGAC 342
DB 523 AGATATTCATCTCAATCTCAATGTCCTACCGAGAGACTCCGACATCTTCAGCCCTACG 582
OY 343 TTTCTGAGGTAAGCAAAATCCCTTCGT-----GTTGAGTGGCAAGCAAGAG 393
DB 583 TGGCTGAGACCCGTGTCGCGCCAGCTGCCACCCCTCACTCTGCGCAGAGC 642
OY 394 AATTTGTGTGCTGGTGTGATTAATCTGAACCTGAGCATGCCAGATCAATTAAC 453
DB 643 GACCTGTGCTGGCGGTGTCACACTGGAAGCCGAGCTACAGCCAGGTGCGCTACTAC 702
OY 454 AATGACTAAGCAAAAGCATTTGAATCCATACCTACGCGCAAGCATTTGGAGATATGTC 513
DB 703 CAGAGCTTCAGAGCTCATCTCAAGGTGAGGTGAGAGCGCTCCACAAAGCCCTGCGCC 762
OY 514 AATGATMAAATTTGATTCCTACATATCTGCTTGAATTTATCTTGAAT 573
DB 763 AAGGGACCA---TGATGAGAGAGCTGTCGCGTACAAATCTTACTGCTTCAAGAAC 819
OY 574 TCAATCCACAGATATACATACCGGAAGAACTAT---ACAATCTTTCTGCTGCTCT 630
DB 820 TCTTGCACCCCGACTACATACACCGAAGCTGTGAGAGAGCCCTGAGGCGCTGCGGCC 879
OY 631 GACCTGTGTCTGTGAGCATCTAGGGAACCTATGAAATTAATTAATTCACAGATTA 690
DB 880 GTGCCCCGTGTGTGAGCCCGCCAGAGAGCAACTACAGAGGTTCCTGACCCAGCAGC 939
OY 691 TTCAATCATGTGAGAGATTAATTAATCTCCAGTACAGTACAAATGATGAGAGAGTC 750
DB 940 TTCAATCATGTGAGAGACTTCCAGAGCCCAAGAGACTGCGCTGACTGAGAGCTG 999
OY 751 GACAAACAAATAAGTTATACCTTACTTACTTACTGAGAGAG 795
DB 1000 GACAAAGACCAAGCCCGCTACCTGAGACTTTCGCTGCGGAG 1044

RESULT 14

US-07-914-281-1
Sequence 1, Application US/07914281
Patent No. 5324663
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/914,281
 FILING DATE: 1992/07/20
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 2363-060-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)521-4500
 TELEFAX: (703)486-2347
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2043 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 ANTI-SENSE: NO
 US-07-914-281-1

Query Match 12.3%; Score 112.2; DB 1; Length 2043;
 Best Local Similarity 51.6%; Pred. No. 3e-24;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

106 TGCATCTCAACGACCGCTTACTGTACACAAATCCATGCGATTGATTCATCAC 165
 343 TGCACATCTACGCGCCGACCAAGGTGTACCCAGGACAGACAGGTCTATGTCACAC 402
 166 CGAGACATCGTGGGA---TGTGACAAATTTACCTCAGCAGAGTATGACACCTTCAG 222
 403 TGGATATCTATGTCACACCTTAAGTACGCTCCACCTTCCCGGACGAGGGGACG 462
 223 AAATGATTTGATGATTTGATGATTCACACACTCACCCTCCCAAGAGTGGCATTTGAG 282
 463 CGGTGATCTGTTTCACTTGGAGCCACCCCTTACTGCGACGACCTGGAAAGCCCTGGAC 522
 283 CACTTGTTAACCTGACTGTGACTTACCGCGCTGATTCAGATATCAAGTCCCTTATGGC 342
 523 AGATATCTCAATCTCAACATGCTCTACCGAGGACTCGACATCTTACGCGCTTACGGC 582
 343 TTCTTGACGGTAAAGCAAAATCCCTTGT-----GTTTGAAGTCCCAAGCAAGAG 393
 583 TGGCTGAGCGGTGTGTCGGGCGCCAGCCGACCCACCGCTCAACCTCTCGGCCAAGACC 642
 394 AAATGATTTGATGATTTGATGATTCACACACTCACCCTCCAGATGCGACATGATTTAC 453
 643 GAGCTGTGCGCTTGGGCGGTGTCACTGGAAGCCGACATCAGCCAGGCTGCGTACTAC 702
 454 AATGAGCTTAAGCAAAAGCATTTGAATCCATACCTAGCGGCAAGCATTTGAGAAATATGTC 513
 703 CAGAGCTGAGGCTCATCTCAAGGTGAGGTGACGAGCGTCCCAAGAGCCCTGCGC 762
 514 AATGATTAATTTGATCTTACATATCTGCTTTGTAATTTTATCTTCTTGTGAATAT 573
 763 AAGGGGACCA---TGATGAGAGCGGTGTCCCGGTACAAATTTTACCTGCGCTTGAAGAC 819
 574 TCAATCCACAAGATTCATCAGGAAAGCTAT---ACAATGCTTTTGTGGCTGGCTCT 630
 820 TCTTTCGACCCCACTACACGAGAACTGTGGAGAAAGCCCTGAGAGCCCTGAGCC 879
 631 GTAACCTGTTTCTGGGACCATCTAGGAAACATATGAGATTAATATTCAGCAGATTTCA 690
 880 GTGCCGTGTGCTGGGCGCCAGAGAAAGCACTAGAGAGGTTCCTGCGCACCCGAGGCC 939
 691 TTCTATCATGTGAAGATTTATTAATCTCCAGATGAGCTAGCAAGATCTGAGAGGAAGTC 750
 940 TTATCTACAGCTGACGACTTCCAGAGCCCAAGGACCTGCGCGCTTGTCCAGAGAGCTG 999
 751 GACAAAACAATAGTATTAATCTTATCTTAACTGAGAGAG 795
 1000 GACAAAGACCAAGCCCGCTACTGAGCTACTTTCGCTGGGAG 1044

RESULT 15
 US-08-393-246-1

; Sequence 1, Application US/08393246
 ; Patent No. 5595900

GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCT

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

ADDRESSEE: P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/393,246

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/220,433

FILING DATE: 30-MAR-1994

APPLICATION NUMBER: US 07/914,281

FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2043 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ANTI-SENSE: NO

US-08-393-246-1

Query Match 12.3%; Score 112.2; DB 1; Length 2043;
 Best Local Similarity 51.6%; Pred. No. 3e-24;

Matches 364; Conservative 0; Mismatches 323; Indels 18; Gaps 4;

106 TGCATCTCAACGACCGCTTACTGTACACAAATCCATGCGATTGATTCATCAC 165
 343 TGCACATCTACGCGCCGACCAAGGTGTACCCAGGACAGACAGGTCTATGTCACAC 402
 166 CGAGACATCGTGGGA---TGTGACAAATTTACCTCAGCAGAGTATGACACCTTCAG 222
 403 TGGATATCTATGTCACACCTTAAGTACGCTCCACCTTCCCGAGGCGCCAGAGGGGACG 462
 223 AAATGATTTGATGATTTGATGATTCACACACTCACCCTCCCAAGAGTGGCATTTGAG 282
 463 CGGTGATCTGTTTCACTTGGAGCCACCCCTTACTGCGACGACCTGGAAAGCCCTGGAC 522
 283 CACTTGTTAACCTGACTGTGACTTACCGCGCTGATTCAGATATCAAGTCCCTTATGGC 342
 523 AGATATCTCAATCTCAACATGCTCTACCGGACGACATCTCCGACATCTTACGCGCTTACGGC 582
 343 TTCTTGACGGTAAAGCAAAATCCCTTGT-----GTTTGAAGTCCCAAGCAAGAG 393

Tue Oct 8 10:18:38 2002

us-09-744-748-5_copy_460_1374.rml

Page 12

```
Db 583 TGGCTGAGACCGTGTCCGCGCCAGCTCCACCCACCGCTCAACCTCTCGGCCAAGAAGC 642
OY 394 AAATGCTGTGCTGGTGTGATGATGAAACCTGTGAGCATGCCAGAGTCAAGTATTAC 453
Db 643 GAGCTGTGTGGCTGTGGCCGTGTCCACGTGAGAGCCGGACTCAGCCAGGTGCGCTACTAC 702
OY 454 AATGAGCTAAGCAAAAGCATTTGAATCCATACCTACGGGCAAGCATTTGAGAAATATGTC 513
Db 703 CAGAGCTGTGAGGCTCATCTCAAGGTGAGCTGTACGAGCTCCCAAAACCCCTGCCC 762
OY 514 AATGATAAAATTTGANTTCTTACCATATCTGCTGTAAATTTATCTTCTCTTGAAT 573
Db 763 AAGGGGACCA--TGATGAGAGCGCTGTCCGCTACAAAGTTCTACCTGCGCTTCGAGAAC 819
OY 574 TCAATCCACAAGATTTACATCAGGAAAAGCTAT--ACAATGCTTTCTGCTGCTCT 630
Db 820 TCCTTGACACCCCGACTACATCAGGAGAGCTGTGAGAGAACCGCCCTGAGAGCCCTGGGCC 879
OY 631 GTACCTGTTGTTGAGACCATCTAGGAAACTATGAGAAATTTATTTCCAGCAGATTCA 690
Db 880 GTGCCCTGTGTGTGGGCCCGAGAGAACTACGAGAGGTCTCCAGCCAGCGCC 939
OY 691 TTCAATTCATGTGGAAGATTATACCTCCCGAGTGAAGTCAAGTATCTGAAGAGATC 750
Db 940 TTCAATCCAGCTGTGAGCATTTCCAGAGCCCGAGAGCTGTGCTGCTGCTGAGAGCTG 999
OY 751 GACAAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 795
Db 1000 GACAGAGACCAAGCCCGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1044
```

Search completed: October 6, 2002, 07:22:03
Job time : 30.6694 secs